AD-A076 024 STANFORD UNIV CALIF DEPT OF PSYCHOLOGY F/G 12/1

PREFERENCE TREES.(U)

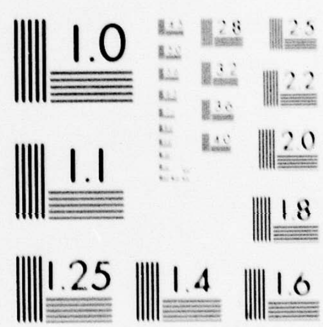
OCT 79 A TVERSKY, S SATTATH N00014-79-C-0077

UNCLASSIFIED NL

| OF |

AD
A076024

END
DATE
FILMED
12-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

PREFERENCE TREES

Amos Tversky

Stanford University

Shmuel Sattath

Hebrew University

October 1979

Preparation of this report was supported by the
Engineering Psychology Programs, Office of Naval Research

ONR Contract N00014-79-C-0077 Work Unit NR197-058

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: (9) Technical Report No. 1, 1 Dec 78 - 1 Oct 79	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) Preference Trees		5. TYPE OF REPORT & PERIOD COVERED: Technical Report Dec. 1, 1978 - Oct. 1, 1979
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Amos Tversky ~~and~~ Shmuel Sattath		8. CONTRACT OR GRANT NUMBER(s): (15) N00014-79-C-0077
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Stanford University, Department of Psychology, Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 197-058
11. CONTROLLING OFFICE NAME AND ADDRESS: Engineering Psychology Programs, Office of Naval Research - Code 455, Arlington, Virgina 22217 (11) 1		12. REPORT DATE: Oct. 1979
		13. NUMBER OF PAGES: 87
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 90		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Making	Preference Tree
Choice Probability	Constrained Choice
Elimination-by-aspects	Agenda Effect

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper develops a probabilistic choice model, called Preference Tree or Pretree, which describes choice as a covert hierarchical elimination process. Each alternative is represented as a measurable collection of aspects and the entire ensemble of aspects is assumed to have a tree structure. At each stage one selects an aspect (i.e., a branch of the tree) with probability that is proportional to its measure and eliminates all the alternatives that do not belong to the selected branch. The process continues

next page

2 to the n th power

cont. until only a single alternative remains. Pretree is considerably more parsimonious than the more general model of Elimination-by-Aspects (EBA) because it has at most 2n rather than 2^n parameters, where n is the number of choice alternatives. At the same time, Pretree is much less restrictive than models (e.g., Luce, 1959) which assume that the strength of preference of x over y is independent of other alternatives. It is shown that the proposed model, which is based on random selection of aspects, is also compatible with a different decision strategy in which the aspects are considered in a fixed sequential order.

The relations between individual and aggregate choice probabilities are discussed, and an additional interpretation of Pretree (and EBA) as an aggregate model is developed. Testable consequences of Pretree are derived, and necessary and sufficient conditions for the existence of tree representations for binary choice probabilities are established. The analysis of several sets of individual and aggregate choice probabilities shows that Pretree fits the observed data and explains the prevalent violations of the constant ratio rule. Tree representations of choice alternatives are constructed using both similarity and preference data. Finally, Pretree is applied to the analysis of choice that is constrained by a partition imposed on the offered set (e.g., an agenda). It is shown that choice probabilities are unaffected by constraints if and only if the constraints are compatible with the structure of the tree. The effect of an agenda on individual choice is investigated experimentally, and its implications to committee decision making and consumer behavior are discussed.

Preference Trees

The analysis of choice behavior has concerned many students of social science. Choices among political candidates, market products, investment plans, transportation modes and professional careers, have been investigated by economists, political scientists and psychologists using a variety of empirical and theoretical methods. An examination of the empirical literature indicates that choice behavior is often inconsistent, hierarchical, and context dependent.

Inconsistency refers to the observation that people sometimes make different choices under seemingly identical conditions. Although inconsistency can be explained as the result of learning, satiation, or change in taste, it tends to persist even when the effects of these factors are controlled or minimized. Furthermore, even in an essentially unique choice situation, which cannot be replicated, people often experience doubt regarding their decisions, and feel that in a different state of mind they might have made a different choice. The observed inconsistency and the experienced uncertainty associated with choice behavior have led several investigators to conceptualize choice as a probabilistic process, and to use the concept of choice probability as a basis for the measurement of strength of preference. (Thurstone, 1927: Luce, 1959; Marschak, 1960).

Choice among many alternatives appears to follow a hierarchical elimination process. When faced with many alternatives (e.g., job offers, houses, cars) people appear to eliminate various subsets of alternatives sequentially according to some hierarchical structure, rather than scanning all the options in an exhaustive manner. This strategy is particularly appealing when the number of alternatives is large and an exhaustive evaluation is either not

feasible or very costly in time and effort. Indeed, these considerations have led several theorists, notably Simon (1957), to modify the classical criterion of maximization, and to view the choice process as a search for an acceptable alternative that satisfies certain criteria. Such a search is naturally executed by a sequential elimination procedure.

Choice behavior appears to be context dependent. That is, the strength of preference of x over y depends on the context of the other available alternatives. Furthermore, choice probability depends not only on the values of the alternatives, but also on their similarity or comparability, see, e.g., Tversky (1972 a). An analysis of the structural relations among the alternatives, therefore, is an essential element of any theory which purports to explain the effects of similarity and context on choice.

The present paper develops a probabilistic, context-dependent choice model--called preference tree--based on a hierarchical elimination process. The first part of the paper illustrates the tree model and investigates its formal properties and their psychological significance. In the second part of the paper, the model is applied to several sets of choice data that are represented as preference trees. The problem of constrained choice is investigated in the third section and the implications of the tree model are discussed in the last section.

THEORY

In order to motivate and develop the theory of preference trees, we discuss first the more general model of elimination by aspects, or EBA. According to this model (Tversky, 1972a, b) each alternative is viewed as a collection of measurable aspects, and choice is described as a covert process of eliminations. At each stage in the process one selects an aspect (from those included in the available alternatives) with probability that is proportional to its measure. The selection of an aspect eliminates all the

alternatives that do not include this aspect, and the process continues until only a single alternative remains. Consider, for example, the choice of a restaurant for dinner. The first aspect selected may be seafood; this eliminates all restaurants that do not serve acceptable seafood. Given the remaining alternatives another aspect, say a price level, is selected and all alternatives that do not meet this criterion are eliminated. The process continues until only one restaurant--that includes all the selected aspects--remains.

In order to characterize this process in formal terms, some notation is introduced. Let $T = \{x,y,z,...\}$ be the total finite set of alternatives under study, and let A,B,C, denote nonempty subsets of T. Let $P(x,A)$ be the probability of choosing alternative x from an offered set A. Naturally $\sum_{x \in A} P(x,A) = 1$ for all $A \subset T$, and $P(x,A) = 0$ for $x \notin A$. For simplicity, we write $P(x,y)$ for $P(x,\{x,y\})$. Choice probabilities are typically estimated from relative frequency of selecting x on repeated choices from A. Next, consider a mapping that associates with each x in T a finite nonempty set $x' = \{\alpha, \beta, ...\}$ of elements which are interpreted as the aspects of x. An alternative x is said to include an aspect α whenever α is an element of x'. The present theory represents choice alternatives as collections of aspects which denote all valued attributes of the options including quantitative attributed (e.g., price, quality) and nominal attributes (e.g., automatic transmission on a car, or fried rice on a menu). The present analysis, however, does not require prior identification of the aspects associated with each alternative.

For any subset A of T, let A' be the set of aspects that belong to at

least one alternative in A, i.e., $A' = \{\alpha | \alpha \varepsilon x' \text{ for some } x \varepsilon A\}$. In particular, T' is the family of all aspects under consideration. For any α in T', let $A_\alpha = \{x \varepsilon A | \alpha \varepsilon x'\}$ denote the set of all alternatives of A that include α. Note that A' is a set of aspects while A_α is a set of alternatives. Using these constructs, the EBA model can now be defined as follows.

A family of choice probabilities $P(x,A), x \varepsilon A \subset T$, satisfies EBA if there exists a non-negative scale u defined on T' such that for all $x \varepsilon A \subset T$

$$(1) \qquad P(x,A) = \frac{\sum_{\alpha \varepsilon x'} u(\alpha) P(x, A_\alpha)}{\sum_{\beta \varepsilon A'} u(\beta)}$$

This recursive formula, which defines the EBA model, expresses the probability of choosing x from A as a weighted sum of the probabilities $P(x,A_\alpha)$ of choosing x from proper subsets of A. It is easy to show that aspects which are common to all the alternatives under consideration do not affect choice probability and can, therefore, be discarded.

Insert Figure 1 here

To illustrate the model, consider the case of three alternatives where $A = \{x,y,z\}$, and let $x' = \{\alpha,\theta,\delta,\lambda\}$, $y' = \{\beta,\theta,\mu,\lambda\}$, and $z' = \{\gamma,\delta,\mu,\lambda\}$, see Figure 1. Thus, $A_\alpha = \{x\}$, $A_\theta = \{x,y\}$, $A_\delta = \{x,z\}$, $A_\lambda = \{x,y,z\}$, etc. Discarding λ which is shared by all alternatives and normalizing the scale u such that $u(\alpha) + u(\beta) + u(\gamma) + u(\delta) + u(\theta) + u(\mu) = 1$ yields

$$P(x,A) = u(\alpha)P(x,A_\alpha) + u(\theta)P(x,A_\theta) + u(\delta)P(x,A_\delta)$$
$$= u(\alpha) + u(\theta)P(x,y) + u(\delta)P(x,z), \text{ where}$$

$$P(x,y) = \frac{u(\alpha) + u(\delta)}{u(\alpha) + u(\beta) + u(\delta) + u(\mu)} = \frac{u(x'-y')}{u(x'-y') + u(y'-x')}$$

This equation for binary choice probabilities coincides with Restle's (1961) model. According to the EBA model, x can be chosen from A (i) if α is selected first, (ii) if θ is selected first and then either α or δ are selected later, (iii) if δ is selected first and then either α or θ are selected later. The probability of choosing x from A, therefore, is the sum of the probabilities associated with these outcomes.

Since there may be many aspects that are unique to x or common to x and y only, α, θ, etc. should be interpreted as collections of aspects. However, for the purposes of the present treatment it is possible to combine, say all the aspects that are unique to x, and treat them as a single aspect. Formally, for any nonempty proper subset A of T let $\bar{A} = \{\alpha | \alpha \varepsilon x' \text{ for all } x \varepsilon A \text{ and } \alpha \not\varepsilon y' \text{ for any } y \varepsilon T-A\}$. Thus, $\bar{A}$ is the set of aspects shared by all alternatives of A that are not shared by any alternative in T-A, and $\{\bar{A} | A \cap T \neq T, \phi\}$ is a partition of the set of all aspects into 2^n-2 aspect sets. To avoid additional notation we use α, β, etc.to denote these aspect sets and supress the distinction between individual aspects and collections of aspects.

If all pairs of distinct alternatives in T are aspect-wise disjoint, i.e., $x' \cap y'$ is null, then $P(x,A_\alpha) = 1$ for any α in x', hence Equation (1) reduces to

$$(2) \qquad P(x,A) = \frac{\sum_{\alpha \varepsilon x'} u(\alpha)}{\sum_{\beta \varepsilon A'} u(\beta)} = \frac{u(x)}{\sum_{y \varepsilon A} u(y)} \qquad \text{where } u(x) = \sum_{\alpha \varepsilon x'} u(\alpha)$$

This is the choice model developed by Luce (1959, 1977). When all choice probabilities are nonzero, Luce's model is equivalent to the assumption that the ratio $P(x,A)/P(y,A)$ is a constant which depends on x and y but not on the offered set A. Hence, it is called the constant-ratio model, abbreviated CRM. This model is simple and parsimonious; it expresses all probabilities of choice among n alternatives in terms of n scale values. (Since the unit of measurement is arbitrary, the number of independent parameters to be estimated is one less the number of scale values). The constant-ratio model, however, fails to account for the effects of similarity between alternatives on choice probability, as shown by several authors, e.g., Debreu (1960), Luce and Suppes (1965), Restle (1961), Rumelhart and Greeno (1971), Tversky (1972 a). The relevant experimental studies were reviewed by Luce (1977).

In contrast, EBA provides a natural explanation of the similarity effect. Furthermore, it has several testable consequences that impose considerable constraints on observed choice probabilities and permit a measurement-free test of a model. The EBA model, however, does not restrict the structure of the aspects in any way, and hence it yields a large number of scale values ($2^n - 2$) which limits its use as a scaling model. In particular, EBA cannot be estimated from binary choice probabilities since the number of parameters exceeds the number of data points. The question arises then whether EBA can be significantly simplified by imposing some structure on the set of aspects. Stated differently, can we formulate an adequate theory of choice that is less restrictive than CRM and more parsimonious than EBA? We can view CRM as the set-theoretical

analogue of a unidimensional representation and EBA as the counterpart of a high dimensional representation. What then is the analog of low dimensionality in a set-theoretical representation?

In this paper we investigate the representation of choice alternatives as a tree-like graph. A graph is a collection of points, called nodes, some of which are linked directly by lines called edges or links. A sequence of adjacent links with no repetitions is called a path. A (rooted) tree is a connected graph without cycles containing a distinguished node called the root. Thus, any two nodes in a tree are joined by a path, and no path starts and ends at the same node. For ease of reference, we place the root at the top of the tree and the terminal nodes at the bottom as in Figure 2. To interpret a rooted tree as a family of aspect sets, we associate each terminal node of the tree with a single alternative in T, and each link of the tree with the set of aspects that are shared by all the alternatives which include (or follow from) that link and are not shared by any of the alternatives which do not include that link. Naturally, the length of each link in the tree represents the measure of the respective set of aspects. Hence, the set of all aspects that belong to a given alternative, is represented by the path from the root of the tree to the terminal node associated with the alternative, and the length of the path represents the overall measure of the alternative.

Insert Figure 2 here

An illustrative example of a tree representation of a menu is presented in Figure 2. The set of alternatives consists of five entrees: steak, roast beef, lamb, sole and trout, that appear as the terminal nodes of the tree. Thus, the link labelled λ represents the aspects shared by all meat entrees but not fish, θ represents the aspects shared by steak and roast beef but not lamb or fish, and γ represents the unique aspects of lamb. The names of the alternatives are displayed vertically and the suggested labels of the clusters (defined by the links) are displayed horizontally.

A tree representation imposes considerable constraints on the family $T^* = \{x' | x \in T\}$ of aspect sets associated with a given set of alternatives. In particular, a tree defines a hierarchical structure on the alternatives in T induced by associating each link α of the tree with the set $T_\alpha = \{x \in T | \alpha \in x'\}$ of all alternatives that include, or follow from, that link. In Figure 2, for example, $T_\mu = \{\text{sole, trout}\}$ and $T_\alpha = \{\text{steak}\}$. It is easy to verify that for any two links α, β in a tree, either $T_\alpha \supset T_\beta$ or $T_\beta \supset T_\alpha$ or $T_\alpha \cap T_\beta$ is empty. The constraints implied by the tree greatly restrict the structure under consideration and drastically reduce the number of parameters from $2^n - 2$ (the number of proper nonempty subjets of T) to $2n - 2$ that corresponds to the maximal number of links in a tree with n terminal nodes. To appreciate the nature of the constraints, note that the paths which connect any three terminal nodes with the root either all meet at the same node, or two paths join at one node while the third path joins them at a higher node, i.e., one that is closer to the root. In Figure 2, for example, 'steak' and 'roast beef' join first and then lamb joins them later.

This property of trees implies the following inclusion rule: for all x,y,z in T, either $x' \cap y' \supset x' \cap z'$ or $x' \cap z' \supset x' \cap y'$. That is, one out of any two binary intersections of three alternatives include the other. Equivalently, any subset of T with three elements contains one alternative, say z, such that $z' \cap x' = z' \cap y'$ which, in turn, is included in $x' \cap y'$. We denote this relation by (x,y)z, with or without a comma. Thus, the tree in Figure 2 is described as ((steak, roast-beef)lamb) (sole, trout). Figure 3a illustrates the inclusion rule by a Venn diagram, and Figure 3b displays the corresponding tree.

Insert Figures 3a and 3b here

A comparison of Figures 1 and 3a reveals that, under the inclusion rule, two out of the three binary intersections coincide with the triple intersection $(x' \cap z' = y' \cap z' = x' \cap y' \cap z')$, hence the number of parameters or aspect sets reduces in this case from 6 to 4, excluding λ that represents the aspects shared by all three alternatives. The following elementary result, proved in the mathematical appendix, shows that the inclusion rule is not only necessary but also sufficient for representation by a tree.

STRUCTURE THEOREM: A family $\{x' | x \epsilon T\}$ of aspect sets is representable by a tree iff either $x' \cap y' \supset x' \cap z'$ or $x' \cap z' \supset x' \cap y'$ for all x,y,z in T.

When the family $\{x' | x \epsilon T\}$ of aspect sets satisfies the inclusion rule, the process of elimination-by-aspects reduces to elimination-by-tree, or EBT for short. That is, one selects a link from the tree (with probability

that is proportional to its length) and then eliminates all the alternatives that do not include the selected link. The same process is then applied to the selected branch, until only one alternative remains. In Figure 3, for example, $P(x,\{x,y,z\}) = u(\alpha) + u(\theta)u(\alpha)/(u(\alpha) + u(\beta))$, and $P(z,\{x,y,z\}) = u(\gamma)$, assuming the measure u is normalized so that $u(\alpha) + u(\beta) + u(\gamma) + u(\theta) = 1$. Elimination by tree, then, is simply the application of elimination by aspects to a tree structure. Note that CRM corresponds to a degenerate tree, or a bush, with only one internal node - the root.

Hierarchical Elimination

The representation of choice alternatives as a tree suggests an alternative decision model in which the tree is viewed as a hierarchy of choice points.[1] This theory, called the hierarchical elimination model or HEM, can be described as follows. One begins at the top of the tree and selects first among the major branches, or the links that follow directly from the root. One then proceeds to the next choice point at the bottom of the selected link, and the process is repeated until the chosen branch contains a single alternative. The probability of choosing an alternative x from an offered set A is the product of the probabilities of selecting the branches containing x at each stage of the process, and the probability of selecting a branch is proportional to its overall weight. For example, the probability of choosing trout from the choice set presented in Figure 2 equals the probability of selecting fish over meat multiplied by the probability of choosing trout over sole. Thus, each node in the tree is treated as a choice point, and one proceeds in order form the top to the bottom of the hierarchy.

To define the hierarchical elimination model in a more formal manner, let A_α denote the set of alternatives in A that include the link α, i.e., $A_\alpha = \{x \epsilon A | \alpha \epsilon x'\}$. Define $\alpha|\beta$ if β follows directly from α, i.e., $A_\alpha \supset A_\beta$, and $A_\gamma \supset A_\beta$ implies $A_\gamma \supset A_\alpha$. Let $u(\alpha)$ denote the length of α, and let $m(\alpha)$ be the measure, or the total length, of all the links that follow from α, including α. In Figure 3b, for example, $\theta|\alpha$, $\theta|\beta$, and $m(\theta) = u(\alpha) + u(\beta) + u(\theta)$. If T* is a tree and $A \subset T$, $A^* = \{x' | x \epsilon A\}$ is also a tree that is referred to as a subtree of T. Naturally, the relation $|$ and the measure u on T* induce corresponding relations and measures on A*. Finally, for $B \subset A$, let $P(B,A)$ denote the probability that the alternative selected from A is also an element of B, i.e., $P(B,A) = \sum_{x \epsilon B} P(x,A)$.

A family of choice probabilities $P(x,A), x \epsilon A \subset T$, is said to satisfy HEM if there exists a tree T*, with a measure u, such that the following three conditions hold

(a) if $\gamma|\beta$ and $\beta|\alpha$ then $P(A_\alpha, A_\gamma) = P(A_\alpha, A_\beta)P(A_\beta, A_\gamma)$

(3) (b) if $\gamma|\beta$ and $\gamma|\alpha$ then $\dfrac{P(A_\alpha, A_\gamma)}{P(A_\beta, A_\gamma)} = \dfrac{m(\alpha)}{m(\beta)}$, provided $P(A_\beta, A_\gamma) \neq 0$.

(c) the above conditions also hold for any subtree A* of T*, with the induced structure on A*.

The first condition implies that the probability of selecting x, say, from T is the product of the probabilities of selecting the branches that contain x at each junction. This condition is readily testable since it is formulated directly in terms of choice probability, with no reference to the scale u. The second condition states that the probabilities of selecting one branch rather than another at a given junction are proportional to the

weights of the respective branches -- defined as the total length of all their links. If we view each junction as a pan balance and the weight of each subtree as mass, then (b) can be interpreted as a weighing process where the probability of choice among subtrees is proportional to their mass. The third condition ensures that (a) and (b) apply not only to the entire tree, but also to any subtree obtained by deleting alternatives from T. Note that the above definition of HEM, like the definition of EBA, excludes in effect the presence of identical alternatives. Thus, we assume that any two alternatives have some distinctive aspects with a nonzero measure, however small.

The notion of hierarchical elimination and the idea of elimination-by-tree represent different conceptions of the choice process that assume a tree structure. EBT describes $P(x,A)$ as a weighted sum of the probabilities $P(x,A_\alpha)$ of selecting x from the various subsets of A. In HEM on the other hand, $P(x,A)$ is expressed as a product of the probabilities $P(A_\alpha,A_\beta)$, $\beta|\alpha$, of selecting a subtree containing x at each level in the hierarchy. Compare, for example, the two formulas for the probability of choosing steak from the set of entrees T displayed in Figure 2. To simplify the notation we suppress the scale u and write α for $u(\alpha)$, etc. Furthermore, the scale is normalized so that $\alpha+\beta+\gamma+\delta+\rho+\theta+\lambda+\mu = 1$. According to EBT, then

$$P(\text{Steak, T}) = \alpha + \theta\left(\frac{\alpha}{\alpha+\beta}\right) + \lambda\left(\frac{\alpha}{\alpha+\beta+\gamma+\theta} + \left(\frac{\theta}{\alpha+\beta+\gamma+\theta}\right) \times \left(\frac{\alpha}{\alpha+\beta}\right)\right)$$

whereas according to HEM

$$P(\text{Steak, T}) = (\alpha+\beta+\gamma+\theta+\lambda) \times \left(\frac{\alpha+\beta+\theta}{\alpha+\beta+\gamma+\theta}\right) \times \left(\frac{\alpha}{\alpha+\beta}\right)$$

The difference in form reflects a difference in processing strategy. EBT assumes free access; that is, each aspect can be selected (as a basis for

elimination) at any stage of the process. On the other hand, HEM assumes sequential access; that is, aspects are considered in a fixed hierarchical fashion. The contrast between models based on random and on sequential access can also be found in theoretical analyses of memory and pattern recognition.

It would appear that EBT is applicable to decisions, such as the selection of a restaurant or the choice of a movie where there is no fixed sequence of choice points, whereas HEM seems appropriate for decisions that induce a natural hierarchy of choice points. A student who has to decide what to do after graduation, for example, is more likely to consider the alternatives in a hierarchical manner. She may first decide whether to go to graduate school, travel, or take a job. And she may not evaluate in detail the available graduate schools, travel plans, or job opportunities, before the initial decision is resolved. The preceding discussion suggests that EBT and HEM capture different decision strategies that might be followed in different situations. However, the following theorem establishes a rather surprising result that, despite the difference in mathematical form and psychological interpretation, the two models are actually equivalent.

EQUIVALENCE THEOREM: EBT and HEM are equivalent. That is, any set of choice probabilities satisfies one model iff it satisfies the other.

The proof of the Equivalence Theorem is given in Section II of the Appendix. It shows that, given a tree T* with a measure u, EBT and HEM yield identical choice probabilities and hence it is impossible to discriminate

between these strategies on the basis of these data alone. It might be possible, however, that other data such as verbal protocols, reaction time or eye movements can be used to distinguish between the two strategies. To avoid confusion, we shall use the term 'preference tree' or 'Pretree' to denote the choice probabilities generated by EBT or by HEM, irrespective of the particular strategy.

An immediate corollary of the equivalence of EBT and HEM is that any alternating strategy consisting of a mixture of EBT and HEM is also equivalent to them. For example, a person may choose a restaurant according to EBT but select an entree according to HEM, or vice versa. It is a remarkable fact that all the various strategies obtained by alternating EBT and HEM yield identical choice probabilities. Thus, Pretree provides a versatile representation of choice that is compatible with both random-access and sequential-access strategies

Consequences

We turn now to discuss general properties and testable consequences of the tree model, starting with the similarity effect. There are two distinct ways in which the similarity between alternatives affect choice probability. First, similarity, or the presence of common aspects creates statistical dependence among alternatives. If x has more in common with y than with z, for example, then the addition of x to the set {z,y} tends to hurt the similar alternative y more than the less similar one z. In the extreme case where x is almost identical to y, the addition of x will divide the probability of choosing y by two while leaving the probability of choosing z unchanged.

Second, similarity facilitates comparison. If x is more similar to y than to z, and $P(y,z) = 1/2$, then $P(x,z)$ will be less extreme than $P(x,y)$, i.e., closer to 1/2. Thus, the more similar pair generally yields a more extreme choice probability because similarity facilitates the comparison between the alternatives.

To illustrate the effects of similarity, consider a hypothetical example of choice among transportation modes. Suppose the available alternatives include two airlines a_1 and a_2, and two trains t_1 and t_2. Suppose further that there is no reason to prefer one airline over the other, but one train t_2 has a very slight but clear advantage over t_1 since it makes one fewer stop along the way. Because the train is more comfortable but the plane is faster suppose one is undecided as to whether to fly or take a train, and hence

$$P(a_1,a_2) = 1/2, \; P(t_2,t_1) = 1, \text{ and } P(a_1,t_1) = P(a_2,t_1) = 1/2.$$

Let $P(x,xyz)$ denote $P(x,\{x,y,z\})$. It follows at once from CRM that $P(t_1,t_1a_1a_2) = 1/3$. Introspection suggests, however, that the selection from $\{t_1,a_1,a_2\}$ is likely to be viewed as a choice between a train and a plane, whence a_1 and a_2 are treated as one alternative that is compared with t_1. Consequently, $P(t_1,t_1a_1a_2)$ will be close to 1/2, while the two other trinary choice probabilities will be close to 1/4. The commonality between a_1 and a_2, therefore, produces a statistical dependence which increases the relative advantage of the odd alternative t_1.

Furthermore, CRM implies that if two alternatives are equivalent in one context, then they are substitutable in any context. That is, it should be possible to substitute one for the other without changing choice probability. Since $P(a_1,t_1) = 1/2$ and $P(t_2,t_1) = 1$, we obtain by substitution $P(t_2,a_1) = 1$. This result, however, seems implausible because the slight

albeit definite advantage of t_2 over t_1 is not likely to eliminate all conflict in the choice between t_2 and a_1. $P(t_2,a_1)$, therefore, is expected to be significantly smaller than one, contrary to CRM. Further discussions of this problem, originally presented by Debreu (1960), can be found in Luce and Suppes (1965, pp. 334-335) and Tversky (1972 a, pp. 282-284).

Insert Figure 4 here

Figure 4 represents the above example as a preference tree. It is easy to verify that, according to the tree model with $\alpha = \beta$ and $\theta + \alpha = \delta$, $P(a_1,a_2) = P(t_1,a_1) = P(t_1,a_2) = 1/2$, $P(t_2,t_1) = 1$, but $P(t_2,a_2) = (\gamma+\delta)/(\gamma+2\delta)$ which approaches 1/2 as γ approaches 0. Furthermore, $P(t_1,t_1a_1a_2) = \delta/(2\delta + \alpha)$ which approaches 1/2 as α approaches 0. Hence the tree model provides a simple and parsimonious account of the similarity effects that are incompatible with CRM.

The effects of similarity on choice probability can also be explained by a Thurstonian or a random utility model such as the additive random aspect model (Tversky, 1972b). In this development each aspect α is represented by a random variable V_α, each x in T is represented by the random variable $V_x = \sum_{\alpha \in x'} V_\alpha$ and, following the random utility model, $P(x,A)$ equals $P(V_x \geq V_y$ for all $y \in A)$. This model, like EBA, accounts for the observed dependence among the alternatives in terms of their common aspects that produce positive correlations among the respective random variables. An additive random aspect model differs from the present development in that the aspects are represented by random variables rather than by constants,

and choice is described as a comparison of sums of random variables rather than as a sequential elimination process. Nevertheless, it was shown (Tversky, 1972b) that EBA, and hence Pretree, is also expressible as a random utility model, though not necessarily an additive one. A random utility analog of the tree model, developed by McFadden (1978), is discussed later.

The following testable properties were derived from EBA (see Tversky 1972a,b; Sattath and Tversky 1976). Since EBT is a special case of EBA, these properties apply to the tree model as well.

Moderate Stochastic Transitivity: If $P(x,y) \geq 1/2$ and $P(y,z) \geq 1/2$ then

$$P(x,z) \geq \min(P(x,y), P(y,z)).$$

This is a probabilistic form of the transitivity assumption. Note that the tree model does not entail the stronger property where 'min' is replaced by 'max'.

Regularity: $P(x,A) \geq P(x,A \cup B)$

The probability of selecting x from a given offered set cannot be increased by enlarging that set.

The Multiplicative Inequality: $P(x,A \cap B) \geq P(x,A)P(x,B)$.

The probability of selecting x from $A \cap B$ is at least as large as the probability of choosing x from both A and B in two independent choices.

The properties discussed so far follow from the general EBA model. We turn now to some new properties of binary choice probabilities that characterize the tree model. To simplify the exposition we introduce the probability ratio $R(x,y) = P(x,y)/P(y,x)$, and restrict the discussion to the case where $P(x,y) \neq 0$ so that $R(x,y)$ is always well-defined. The results can be readily extended to deal with choice probabilities that equal 0 or 1.

Consider first the case of three alternatives, and note that any subtree of three elements has the form portrayed in Figure 5, except for the permutation of the alternatives and the possibility of vanishing links. We use the parentheses notation to describe the structure of the tree , e.g., the tree in Figure 5 is described by $(xy)z$ and the tree in Figure 4 by $(a_1a_2)(t_1t_2)$.

Insert Figure 5 here

Using the notation of Figure 5 it follows at once that $R(x,y) = \alpha/\beta$ is more extreme (i.e., further from one) than $R(x,z)/R(y,z) = (\alpha+\theta)/(\beta+\theta)$. Hence any three elements, that form a subtree $(xy)z$, satisfy the following trinary condition.

(4) If $R(x,y) \geq 1$ then $R(x,y) \geq \dfrac{R(x,z)}{R(y,z)} \geq 1$,

where a strict inequality in the hypothesis implies strict inequalities in the conclusion, and an equality in the hypothesis implies equalities in the conclusion.

The trinary condition (4) reflects the similarity hypothesis in that the commonality between alternatives enhances their discriminability. This is seen most clearly in the case where $\theta > 0$, $\alpha > \beta$, and $\beta + \theta = \gamma$, i.e., $R(x,y) > 1$ and $R(y,z) = 1$, see Figure 5. According to the trinary condition $R(x,y) = \alpha/\beta > (\alpha+\theta)/(\beta+\theta) = R(x,z)$. Although y and z are pair-wise equivalent, $P(x,y)$ exceeds $P(x,z)$ because x shares more aspects with y

than with z. Note that when θ vanishes, $R(x,y) = R(x,z)/R(y,z)$ as required by CRM. In this case, where (xy)z, (xz)y and (zy)x all hold we omit the parentheses altogether and write xyz.

Next, let us consider sets of four alternatives. It is easy to verify that, up to permutations of alternatives, any subtree of four elements has one of the two forms displayed in Figure 6, including degenerate forms with one or more vanishing links.

Insert Figure 6 here

It follows readily that in the tree (xy)(vw) portrayed in Figure 6a

$$(5)\qquad \frac{R(x,v)}{R(y,v)} = \frac{(\alpha+\theta)/(\gamma+\lambda)}{(\beta+\theta)/(\gamma+\lambda)} = \frac{(\alpha+\theta)/(\delta+\lambda)}{(\beta+\theta)/(\delta+\lambda)} = \frac{R(x,w)}{R(y,w)}$$

If we interpret $R(x,v)/R(y,v)$ as an indirect measure of preference for x over y, measured relative to a standard v, then the above quarternary condition asserts that this measure is the same for different standards (v and w) provided the pairs (x,y) and (v,w) belong to distinct clusters.

If the relation among the four alternatives under consideration has the form depicted in Figure 6b, that is ((xy)v)w, then the following quarternary condition holds.

$$(6)\qquad \frac{R(x,v) - R(y,v)}{R(x,w) - R(y,w)} = \frac{(\alpha-\beta)/\gamma}{(\alpha-\beta)/\delta} = \frac{(\alpha+\theta-\gamma)/\gamma}{(\alpha+\theta-\gamma)/\delta} = \frac{R(x,v) - R(v,v)}{R(x,w) - R(v,w)}$$

Note that under CRM the quarternary conditions hold for any four alternatives.

At this point, the reader may suspect that the consideration of more elaborate tree structures involving larger sets of alternatives will yield additional independent consequences. However, the following theorem shows that the trinary and the quarternary conditions are not only necessary but they are also sufficient to ensure the representation of binary choice probabilities as a preference tree.

REPRESENTATION THEOREM: A set of nonzero binary choice probabilities satisfies the tree model with a given structure iff the trinary (4) and the quarternary (5 & 6) conditions are satisfied relative to that structure.

The theorem shows that if Equations (4), (5) and (6) are satisfied relative to some tree structure, then there exists a ratio scale u defined on that structure such that

$$P(x,y) = \frac{u(x'-y')}{u(x'-y') + u(y'-x')} \quad \text{or} \quad R(x,y) = \frac{u(x'-y')}{u(y'-x')}$$

Recall that $u(x'-y')$ is the measure of the aspects of x that are not included in y, or the length of the path from the terminal node associated with x to the meeting point of the paths from x and y to the root.

The proof of the Representation Theorem is presented in Section III of the Appendix. This result shows, in effect, how to construct a preference tree from binary choice probabilities whenever the necessary conditions hold. The trinary and quarternary conditions are readily testable--given any specified tree structure. Moreover they can be used to determine which structure, if any, is compatible with the data. Recall that at least one permutation of every triple must satisfy Equation (4), and at least one

permutation of every quadruple must satisfy Equation (5) or (6). Hence, by finding the appropriate permutations of all triples and quadruples, any tree structure that is compatible with the data will emerge. It is readily verified that the scale values (i.e., the length of the links associated with a particular tree structure) are uniquely determined up to an arbitrary unit of measurement, except when all binary choice probabilities are one-half. The tree structure, however, is not always unique. That is, a given set of binary choice probabilities could be compatible with more than one tree structure. An example of this kind is presented in Section IV of the Appendix along with a proof of the proposition that the tree structure is uniquely determined by the set of binary and trinary choice probabilities.

Furthermore, if both binary and trinary choice probabilities are available, they must satisfy the following conditons. Suppose the tree model holds with (xy)z, see Figure 5, then

$$(7)\qquad \frac{P(x,z)}{P(z,x)} = \frac{\alpha+\theta}{\gamma} \geq \frac{\alpha+\theta\alpha/(\alpha+\beta)}{\gamma} = \frac{P(x,xyz)}{P(z,xyz)} \quad \text{and}$$

$$(8)\qquad \frac{P(x,y)}{P(y,x)} = \frac{\alpha}{\beta} = \frac{\alpha+\theta\alpha/(\alpha+\beta)}{\beta+\theta\beta/(\alpha+\beta)} = \frac{P(x,xyz)}{P(y,xyz)}$$

provided all choice probabilities are nonzero Thus, according to the tree model with (xy)z, the constant-ratio rule (8) holds for the adjacent pair (x,y) but not for the split pair (x,z). Note that this rule is violated by (7) in the direction implied by the similarity hypothesis for

(xy)z. Since y is closer to x than to z in that structure (in the sense that $y' \cap x' \supset y' \cap z'$), the addition of y to the set {x,z} reduces the probability of choosing x proportionally more than the probability of choosing z. On the other hand, since z is equally distant from x and from y (in the sense that $x' \cap z' = y' \cap z'$) the addition of z to the set {x,y} reduces the probabilities of choosing x and y by the same factor.

Aggregate Probabilities

So far, we have modeled the process by which an individual chooses among alternatives. Because of the difficulties in obtaining independent repeated choices from the same individual, most available data consist of the proportions of individuals who selected the various alternatives, referred to as group data or aggregate probabilities. It should be emphasized that these data do not pertain to group decision making, they merely characterize the aggregate preferences of different individuals.

It is well-known that most probabilistic models for individual choice (including CRM and EBA) are not preserved by aggregation. That is, group probabilities could violate the model even though each individual satisfies it, and vice versa. Consider, for instance, the case of three individuals 1, 2, 3 and three alternatives x, y, z. Suppose the observed choice probabilities P(x,y), P(y,z) and P(z,x) are, respectively, .75, .75 and .15 for individual 1; .15, .75 and .75 for individual 2; and .75, .15 and .75 for individual 3. The individual choice probabilities all satisfy EBA, but the expected aggregate probabilities .55, .55 and .55, respectively, violate EBA. Hence, the validity of EBA as a model for individual choice is neither necessary nor sufficient for its validity as an aggregate model. Nevertheless, we contend that similar

principles govern both types of choice data, and propose a new interpretation of EBA as an aggregate model.

Suppose each individual chooses in accord with the following sequential elimination rule. Given an offered set A, select some (nonempty) subset of A, say B, and eliminate all the alternatives that do not belong to B. Repeat the process until the selected subset consists of a single alternative. Let $Q_A(B)$ be the proportion of subjects who first select B when presented with the offered set A, i.e., the proportion of subjects who eliminate all elements of A-B in the first stage. Naturally, $\sum_{B_i \subset A} Q_A(B_i) = 1$, and $Q_A(A) = 1$ iff A consists of a single alternative. Note that $Q_A(B)$ is an elimination probability--not a choice probability. The two constructs are related via the following equation.

(9) $$P(x,A) = \sum_{B_i \subset A} Q_A(B_i)P(x,B_i).$$

Thus, the proportion of subjects who choose x from A is obtained by summing, over all proper subsets B_i of A, the proportion of individuals who first select B_i multiplied by the proportion of subjects who choose x from the selected subset. This general elimination model, by itself, does not restrict the observed choice probabilities because we can always set $Q_A(B) = P(x,A)$ if $B = \{x\}$, and $Q_A(B) = 0$ otherwise. Nevertheless, it provides a method for characterizing probabilistic choice models in terms of the constraints they imposed on the elimination probabilities.

A family of elimination probabilities, $Q_A(B)$, $B \subset A \subset T$, satisfies proportionality iff for all A, B, C, B_i, C_j in T,

$$(10) \quad \frac{Q_A(B)}{Q_A(C)} = \frac{\Sigma Q_T(B_i)}{\Sigma Q_T(C_j)}$$

where the summations range, respectively, over all subsets B_i,C_j of T such that $B_i \cap A = B$ and $C_j \cap A = C$. It is assumed that the denominators are either both positive or both zero. This condition implies that, for any $A \subset T$, the values of Q_A are computable from the values of Q_T. More specifically, the percentage of subjects who first select B, when presented with the offered set A, is proportional to the percentage of subjects, presented with the total set T, who first select any subset B_i *that includes in addition to* B only elements that do not belong to A.

To illustrate the proportionality condition, consider the choice among entrees. Let T = {r,s,t} and A = {r,t}, where r, s and t denote, respectively, roast beef, steak and trout. According to proportionality, therefore,

$$\frac{Q_A(r)}{Q_A(t)} = \frac{Q_T(r) + Q_T(r,s)}{Q_T(t) + Q_T(t,s)}$$

Note that in the binary case, where A = {r,t}, $Q_A(r) = P(r,A) = P(r,t)$.

The rationale behind the proportionality condition is the assumption that, upon restricting the offered set from T to A, all individuals who first selected $B \cup C$ from T, $C \subset T-A$, will now select B from A since the alternatives of C are no longer available. For example, those who first

selected {r,s} from T will select roast beef when restricted to A because now steak is not on the menu. The following theorem shows that the (aggregate) process described above is compatible with EBA.

AGGREGATION THEOREM: A set of aggregate choice probabilities on T are compatible with EBA iff there exist elimination probabilities on T that satisfy Equations (9) and (10).

The proof of this theorem is readily reduced to earlier results, see the Appendix in Tversky (1972a) and Theorem 2 in Tversky (1972b). It shows that if (9) and (10) hold then

$$P(x,A) = \frac{\Sigma Q(B_i)P(x,A \cap B_i)}{\Sigma Q(B_i)}$$

where $Q(B_i) = Q_T(B_i)$, and the summations range over all $B_i \subset T$ such that $B_i \cap A$ is nonempty. This form, in turn, is shown to be equivalent to EBA. Hence, the Aggregation Theorem provides a new interpretation of EBA as a model for group data.

It is instructive to compare the above version of the EBA model to the original version defined in Equation (1). First, note that the scale Q(B) is not a measure of the overall value of the alternatives of B. Rather, it reflects the degree to which they form a good cluster, as evinced by the proportion of subjects who first selected B when presented with T. The counterpart of Q(B) in the original version of the EBA model is $u(\bar{B})$, the measure of the aspects that belong to <u>all</u> alternatives of B, and do not belong to <u>any</u> alternative in T-B.

The individual version of the EBA model assumes that at any point in time one has a fixed ordering of the relevant aspect-sets which, in turn, induces a (lexicographic) ordering of the available alternatives. However, at a different point in time, one may be in a different state of mind which yields different ordering of aspects and alternatives. Indeed, the stochastic component was introduced into the model to accommodate such momentary fluctuations. The new aggregate version of EBA assumes that each individual has a fixed ordering of the relevant aspect-sets, and the stochastic component of the model is associated with differences between individuals rather than with changes within an individual. Hence, the former version explains choice probabilities in terms of an intra-individual distribution of states of mind, whereas the latter version explains the data in terms of an inter-individual distribution of tastes.

The EBA model may provide a useful model of aggregate data because the same principles that give rise to EBA as a model of individual choice appear to apply to group data. As a case in point, let us reexamine the similarity effect using the transportation problem discussed earlier. Suppose the group is divided equally between the train t_1 and the plane a_1, and is also equally divided between the two airlines a_1 and a_2. Hence,

$$P(t_1,a_1) = P(a_1,a_2) = 1/2$$

We propose that the proportion of individuals who choose the train t_1 from the offered set $\{t_1,a_1,a_2\}$ lies between 1/2 and 1/3 because the addition of a_2 to $\{t_1,a_1\}$ is likely to affect those who chose a_1 more than those who chose t_1. More generally, the addition of a new alternative or product (e.g., a low-tar cigarette or a liberal candidate) hurts similar alternatives (e.g., other low-tar

cigarettes, and liberal candidates) more than less similar alternatives.

Furthermore, as in the case of individual choice, the similarity between options appears to enhance the discrimination between them. Suppose that each individual prefers train t_2 over train t_1 since it is slightly faster. Suppose further that the group is equally divided between a_1 and t_1, so that $P(a_1,t_1) = 1/2$. Contrary to CRM which implies $P(t_2,a_1) = 1$, we predict that $P(t_2,a_1)$ is likely to be between 1/2 and 1 because many of those who prefer a_1 over t_1 are not likely to switch from a plane to a train because of the slight, albeit clear, advantage of the faster train. Since the same correlational pattern emerges from both individual and group data, the EBA model may be applicable to both, although the assumptions and the parameters of the model have different interpretations in the two cases.

Consider, for example, the assumption that the alternative set $T = \{a_1,a_2,t_1\}$ in the transportation problem has a tree structure $(a_1\ a_2)t_1$. In the individual version, the tree assumption implies that any aspect that is shared by the train and any one of the airlines is also shared by the other airline. In the aggregate case, the tree assumption entails that both $Q_T(a_1,t_1)$ and $Q_T(a_2,t_1)$ vanish, that is, nobody eliminates from T one airline only. Hence, if all individuals share the same tree structure but not necessarily the same preferences, the aggregate data will generally exhibit the same qualitative structure. The actual measure, derived from aggregate data however, does not relate to the measures derived from individual data in any simple manner.

APPLICATIONS

In this section we apply the tree model to several sets of individual and aggregate choice probabilities reported in the literature, construct tree representations for these data and test Pretree against CRM. As was demonstrated in the previous section, the trinary and the quarternary conditions provide necessary and sufficient conditions for the representation of binary choice probabilities as a preference tree. For error-free data, therefore, these conditions can be readily applied to find a tree structure that is compatible with the data. Since data are fallible, however, the construction of the most appropriate tree structure, the estimation of link-lengths and the evaluation of the adequacy of the tree model, pose non-trivial computational and statistical problems.

In the present paper, we do not develop a comprehensive solution to the construction, estimation, and evaluation problems. Instead, we rely on independent judgments (e.g., similarity data) for the construction of the tree, and employ standard iterative maximization methods to estimate its parameters. To evaluate goodness-of-fit we test the tree model assuming the hypothesized tree structure, against the binary version of Luce's constant-ratio model.

It has been shown by Luce (1959) that the binary CRM, according to which $P(x,y) = v(x)/(v(x)+v(y))$, is essentially equivalent to the following product rule

(11) $P(x,y)P(y,z)P(z,x) = P(x,z)P(z,y)P(y,x)$, i.e., $R(x,y)R(y,z)R(z,x) = 1$

Thus, any two intransitive cycles through the same set of alternatives are equiprobable. On the other hand, the trinary condition (4) yields

(12) If $P(x,y) > \frac{1}{2}$ and $(xy)z$ then $R(x,y)R(y,z)R(x,z) > 1$,

or $P(x,y)P(y,z)P(z,x) > P(x,z)P(z,y)P(y,x)$.

Any hypothesized tree structure, therefore, can be examined to test whether the product rule is violated in the predicted direction.

The analysis of the data proceeds as follows. We start with a given set of individual or collective pair comparison data along with a hypothesized tree structure, derived from a priori considerations or inferred from other data. Maximum likelihood estimates for both CRM and Pretree are obtained using Chandler's (1969) iterative program (STEPIT), and the two models are compared via a likelihood ratio test. In addition, we perform an estimate-free comparison of the two models, by contrasting the product rule (11) and the trinary inequality (12).

Choice between Celebrities

Rumelhart and Greeno (1971) investigated the effects of similarity on choice probability, and compared the choice models of Luce (1959) and Restle (1961). The stimuli were 9 celebrities including three politicians (L. B. Johnson, Harold Wilson, Charles DeGaulle), three athletes (Johnny Unitas, Carl Yastrzemski, A. J. Foyt), and three movie stars (Brigitte Bardot, Elizabeth Taylor, Sophia Loren). The subjects (N=234) were presented with all 36 pairs of names and were instructed to choose for each pair "the person with whom they would rather spend an hour discussing a topic of their choosing".

On the basis of a χ^2 test for goodness-of-fit, applied to the aggregate choice probabilities, Rumelhart and Greeno (1971) were able to reject Luce's model ($\chi^2(28) = 78.2$, $p < .001$) but not a particular version of Restle's model ($\chi^2(19) = 21.9$, $p > .25$). Recall that Restle's model coincides with the binary form of the EBA model.

The list of celebrities used in this study naturally suggests the following tree structure with three branches corresponding to the three different occupations represented in the list: (LBJ, HW, CDG) (JU, CY, AJF) (BB, ET,SL). The estimates of the parameters of the tree,[2] displayed in Figure 7, are identical to those obtained by Edgell, Geisler and Zinnes (1973), who corrected the procedure used by Rumelhart and Greeno (1971) and proposed a simplification of the model which amounts to the above tree structure. The tree model appears to fit the data quite well ($\chi^2(25) = 30.0$, $p > .20$), although it has only three more parameters than Luce's model.

Insert Figure 7 here

Since Pretree includes CRM, the likelihood-ratio test can be used to test and compare them. The test is based on the fact that if Model 1 is valid and includes Model 2 then, under the standard assumptions, $-2\ln(L_1/L_2)$ has a χ^2 distribution with d_1-d_2 degrees of freedom, where L_1 and L_2 denote the likelihood functions of models 1 and 2, while d_1 and d_2 denote the respective numbers of parameters. If the inclusive model is saturated, i.e., imposes no constraints, then the above test is equivalent to the common χ^2 test for goodness of fit. When the likelihood-ratio test is applied to the present data, CRM is rejected in favor of Pretree, $\chi^2(3) = 48.2$,$p < .001$. The average absolute deviation between predicted and observed probabilities is .036 for CRM and .023 for Pretree.

It should be noted (see Falmagne, Reference Note 1, 1979) that the test statistics for Pretree does not have an exact χ^2 distribution because the parameter space associated with the model is constrained not only by

the equations implied by the quarternary conditions, but also by the trinary inequality. The result, however, is a stricter test of Pretree since the inequalities imposed on the solution can only reduce goodness of fit.

Since the product rule (11) and the trinary inequality (12) are the key binary properties that give rise, respectively, to CRM and Pretree, it is instructive to compare them directly. Using the tree structure presented in Figure 7, the trinary inequality applies in 9 x 6 = 54 triples and it is satisfied in 89% of the cases. Because the various triples are not independent, no simple statistical test is readily available. To obtain some indication about the size of the effect, we computed the value of $R(xyz) = R(x,y)R(y,z)R(z,x)$ for all triples satisfying $(xy)z$ and $R(x,y) > 1$. The median of these values equals 1.40, and the interquartile range is (1.13, 1.68). Recall that under CRM the trinary inequality is expected to hold in 50% of the cases, and the median $R(xyz)$ should equal one. The summary statistics for all the studies in this section, are presented in Table 1.

Political Choice

The next three data sets were obtained from Lennart Sjöberg, who collected both similarity and preference data for several sets of stimuli, and showed a positive correlation between interstimulus distances (derived from multidimensional scaling) and the standard deviation of utility differences (derived from a Thurstonian model). Sjöberg (1977) and Sjöberg and Capozza (1975) conducted two parallel studies of preferences for Swedish and Italian political parties. In these experiments, 215 Swedish students and 195 Italian students were presented with all pairs of the seven leading Swedish

and Italian parties, respectively. The subjects first rated the similarity between all 21 pairs of parties on a scale from 1 to 9, and then indicated for each pair which party they prefer. In addition, the subjects were presented with all 35 triples of parties and asked to choose one party from each triple.

The average similarities between the parties were first used to construct an additive similarity tree according to the ADDTREE method developed by Sattath and Tversky (1977). In this construction, which generalizes the familiar hierarchical clustering scheme, the stimuli are represented as terminal nodes in a tree so that the dissimilarity between stimuli corresponds to the length of the path that joins them. For illustration, we present in Figure 8 the additive tree (ADDTREE) solution for the similarities between the Swedish parties. The product-moment correlation between rated similarities and path-length is -.96. Assuming the tree structure derived from ADDTREE, Chandler's (1969) STEPIT program was employed to search for maximum likelihood estimates of the parameters of Pretree--using the observed choice probabilities. The obtained preference tree for the Swedish data is presented in Figures 9, and the preference tree for the Italian data is presented in Figure 10.

Insert Figures 8, 9, 10 here

Several comments about the relations between similarity and preference trees are in order. First, the rules for computing dissimilarity and preference from a given tree are quite different. The dissimilarity between x and y

is represented by the length of the path (i.e., the sum of the links) that connects x and y, while the degree of preference R(x,y) is represented by the ratio of the respective paths. Second, the numerical estimates of the links in the two representations tend to differ systematically. In general, the distances between the root and the terminal nodes vary much more in a preference tree (due to the presence of extreme choice probabilities) than in a similarity tree. Furthermore, some links that appear in the similarity tree sometimes vanish in the estimation of Pretree (as can be seen by comparing Figures 8 and 9) indicating the presence of aspects that affect judged similarity, but not choice probability. Third, the root in a similarity tree is essentially arbitrary since the distance between nodes is unaffected by the choice of root. The probability of choice in Pretree, however, is highly sensitive to the choice of a root. Consequently, several alternative roots were tried and the best-fitting structure was selected in each case.

Tests of goodness of fit indicate that Pretree provides an excellent account of the Swedish data $\chi^2(11) = 5.8$, $p > .5$, with an average absolute deviation of .012, compared with $\chi^2(15) = 49.1$, $p < .001$, with an average absolute deviation of .038 for CRM. Pretree also provides a reasonable account of the Italian data $\chi^2(11) = 19.5$, $p > .05$, with an average absolute deviation of .023, compared with $\chi^2(15) = 67.6$, $p < .001$, with an average absolute deviation of .042 for CRM. The applications of the likelihood ratio test indicate that Pretree fits these data significantly better than CRM; the test statistics are $\chi^2(4) = 43.3$, $p < .001$, for the Swedish data and

$\chi^2(4) = 48.1$, $p < .001$, for the Italian data. Furthermore, for the Swedish data, the trinary inequality is satisfied in 96% of the cases (N = 23), the median R(xyz) equals 1.73, and the interquartile range is (1.38, 2.27). For the Italian data, the trinary inequality is satisfied in 78% of the cases (N = 18), the median R(xyz) equals 1.74, and the interquartile range is (.93, 2.78).

The availability of both binary and trinary probabilities in the political studies permitted an additional test of Pretree. Recall from (7) that the tree model implies

$$\frac{P(x,z)}{P(z,x)} > \frac{P(x,xyz)}{P(z,xyz)} \qquad \text{provided } (xy)z,$$

while CRM implies that the two ratios are equal. For the Swedish data, the above inequality is satisfied in 87% of the cases (N = 46), the median $P(x,z)P(z,xyz)/P(z,x)P(x,xyz)$ equals 1.28, and the interquartile range is (1.12, 1.64). For the Italian data, the inequality is satisfied in 81% of the cases (N = 36), the median of the above product ratio equals 1.19, and the interquartile range is (.86, 2.28). Note that under CRM

$$P(x,z)P(z,xyz)/P(z,x)P(x,xyz) = u(x)u(z)/u(z)u(x) = 1.$$

Choice between Academic Disciplines

In a third study conducted by Sjöberg (1977), the alternatives consisted of the following twelve academic disciplines that comprise the social science program at the University of Göteborg: Psychology, Education, Sociology, Anthropology, Geography, Political Science, Law, Economic History Economics, Business Administration, Statistics, Computer Science. A group of 85 students from that university first rated the similarity between all pairs of disciplines on a 9 point scale, and then indicated for each

of the 66 pairs the discipline they prefer.

As in the two preceding analyses, the tree structure was obtained via ADDTREE, and STEPIT was employed to search for maximum likelihood estimates of the parameters. The resulting preference tree for the choice between the twelve social sciences is presented in Figure 11.

Insert Figure 11 here

A χ^2 test for goodness of fit yields $\chi^2(50) = 45.5$, $p > .25$ for Pretree, compared with $\chi^2(55) = 69.1$, $p > .05$, for CRM, and the likelihood ratio test rejects CRM in favor of Pretree, $\chi^2(5) = 23.6$, $p < .001$. The average absolute deviation between predicted and observed probabilities is .025 for Preetree and .035 for CRM. Finally, the trinary inequality is satisfied in 84% of the cases (N = 86), the median $R(xyz)$ equals 1.52, and the interquartile range is (1.21, 1.86).

Choice Between Shades of Gray

In a classic study of unfolding theory, Coombs (1958) used as stimuli 12 patches of grey that vary in brightness. The subjects were presented with all possible sets of 4 stimuli, and were asked to rank them from the most to the least representative grey. Binary choice probabilities were estimated for each subject by the proportion of rank-orders in which one stimulus was ranked above the other. The data provided strong support for Coomb's probabilistic unfolding model in which the stimuli are represented as random variables, and the derived choice probabilities reflect momentary fluctuations in one's perceptions of the stimuli as well as in one's notion of the ideal gray.

Insert Figure 12 here

To represent Coombs' data as a tree, consider a line representing variation in brightness (with white and black at the two endpoints) that is folded in the middle at a point corresponding to the prototypical gray. The stimuli can now be represented as small branches stemming from this folded line, see Figure 12. Because of the large number of zeros and ones in these data,we did not attempt to estimate the tree. Instead, we inferred the characteristic folding point of each subject from the data and used the induced tree structure to compare, separately for each subject, the trinary inequality against the product rule, letting $P(x,y)$ denote the probability that x is judged to be farther than y from the prototypical gray. Triples involving zero probability were excluded from the analysis. The results for each one of the four subjects, presented in the bottom part of Table 1, show that the product rule (11) is violated in the manner implied by the trinary inequality (12)

Insert Table 1 here

Table 1 summarizes the analyses of the studies discussed in this section. The left-hand part of the table describes the statistics for the trinary inequality, where N is the number of tested triples, π is the percentage of triples that confirm the trinary inequality, R is the median value of $R(xyz) = R(x,y)R(y,z)R(z,x)$, while R_1 and R_3 are the first and third quartiles of the distribution of $R(xyz)$. The right-hand part of Table 1 describes

the measures of goodness of fit for both CRM and Pretree, where d is the average absolute deviation between observed and predicted choice probabilities.

Tree Representation of Choice Data

The examination of the trinary inequality provides an estimate-free comparison of CRM and Pretree. The results described in Table 1 show that, in all data sets, CRM is violated in the direction implied by the similarity hypothesis and the assumed tree structure. The statistical tests for the correspondence between models and data indicate that Pretree offers an adequate account of the data that is significantly better than the account offered by CRM. Apparently, the introduction of a few additional parameters, that correspond to aspects shared by some of the alternatives, results in a substantial improvement in goodness of fit. Furthermore, Pretree yields interpretable hierarchical representations of the alternatives under study along with the measures of the relevant aspect sets.

The preceding analyses relied on similarity data or on a priori considerations to construct the tree structure, and used choice probabilities to test the model and to estimate the tree. This procedure avoids the difficulty involved in using the same data for constructing the tree and for testing its validity. It is also attractive because similarity data are easily obtained, and because they are typically more stable and less variable than preferences. An examination of Sjöberg's data, for example, shows that subjects who reveal markedly different preferences tend, nevertheless, to exhibit considerable agreement in judgments of similarity. The only drawback of this procedure is that it fails to produce the best tree whenever the similarities and the preferences follow different structures. The

development of an effective algorithm for constructing a tree from fallible preferences and the development of appropriate estimation and testing procedures remain open problems for future research.

The correspondence between the observed and the predicted choice probabilities indicate that the tree structures inferred from judgments of similarity generally agree with the structures implied by the observed choice probabilities. This result supports the notion of correspondence between similarity and preference structures, originated by Coombs (1964), and underscores the potential use of similarity scaling techniques in the analysis of choice behavior. Other analyses of the relations between the representations of similarity and of preference, based on multidimensional scaling, are reported in Carroll (1972), Nygren and Jones (1977). Sjöberg (1977) and Stefflre (1972).

CONSTRAINED CHOICE AND THE EFFECT OF AGENDA

The preceding development, like other models of choice, deals with the selection of a single element from some offered set. The present section investigates choice that is constrained by a partition imposed on the offered set. For example, the choice of an alternative from the set {x,y,v,w} can be constrained by the requirement to choose first between {x,y} and{v,w} and then to choose a single element from the selected pair. Constraints of this type are quite common: they could be imposed by others, induced by circumstances,or adopted for convenience.

For example, the decision regarding a new appointment is sometimes introduced as an initial decision between a senior or a junior appointment, followed by a later choice among the respective junior or senior candidates. Deadlines and other time limits provide another source of constraint. Suppose the alternatives of $A \subset T$, for example, are no longer available after April 1st. Prior to this date, therefore, one has to decide whether to choose an element of A, or to select an element from T - A,in which case the choice of a particular element can be delayed. The selection of an agenda and the grouping of options for voting (which have long been recognized as influential procedures) are familiar examples of external constraints.

There are many situations, however, in which a person constrains his choice to reduce cost or effort. Consider, for example, a consumer who intends to purchase one item from a set {x,y,v,w} of 4 competing products. Suppose there are two stores in town that are quite distant from each other; one store carries only x and y, while the other carries only v and w. Under

such circumstances, the consumer is likely to select first a store and then a product, because he has to decide which store to enter but he does not have to choose a product before entering the store. Similarly, people typically select a restaurant first and an entree later -- even when they are thoroughly familiar with the available menus. Thus, the need to make some decisions (e.g., of a restaurant)at an early stage and the common tendency to delay decisions (e.g., of an entree) to a later stage constrain the sequence of choices leading to the selected alternative.

The effect of an agenda on group decision making has been investigated by an economist, Charles R. Plott, and a lawyer, Michael E. Levine, from Caltech. Levine and Plott (1977) conducted an ingenious study of a flying club, to which they belong, whose members had to decide on the size and composition of the club's aircraft fleet. There were a few hundred competing alternatives, and the group was to meet once and decide by a majority vote. Levine and Plott constructed an agenda designed to maximize the chances of selecting the alternative they preferred. The group followed this agenda, and, indeed, chose the option favored by the authors. A second study demonstrated the impact of agenda under controlled laboratory conditions. Plott and Levine (1978) developed a model for individual voting behavior and used it to construct for each alternative an agenda for the group, designed to enhance the selection of that alternative. The results indicate that, although the specific model was not fully supported, the imposed agenda had a substantial effect on group choice.

A Theoretical Analysis

An agenda or a constraint imposed on an offered set imposes a hierarchical structure or a tree on that set. Suppose, for example, that {B,C,D} is a partition of A; hence, under the constraint [[B][C]][D] the choice of an alternative from A proceeds by first choosing between D and $B \cup C$ and then choosing between B and C--if D is eliminated in the first stage. It is essential to distinguish here between the intrinsic tree structure (defined in terms of the relations among the aspects that characterize the alternatives) and the imposed structure that characterizes the external constraints. The choice among {x,y,v,w}, for example, whose aspects form the tree (xy)(vw) may be constrained by the requirement to choose first between {x,w} and {y,v}. To avoid confusion we use parentheses, e.g., (xy)v, to characterize the intrinsic tree, and brackets, e.g., [xy]z, to denote the imposed constraints.

Let $P(x,[A][B])$, $x \varepsilon A$, $A \cap B = \phi$, denote the probability of selecting x from $A \cup B$ subject to the constraint of choosing first between A and B. The present treatment is based on the following assumption.

$$(13) \quad P(x,[A][B]) = P(x,A)P(A,A \cup B) = P(x,A) \sum_{y \varepsilon A} P(y,A \cup B).$$

That is, the probability of choosing x under [A][B] is decomposable into two independent choices: the choice of x from A, and the choice of A from [A][B]. Furthermore, the latter choice is reduced to the selection of any element of A from the offered set $A \cup B$. Hence for A = {x,y} and B = {v,w}, $P(x,[xy][vw]) = P(x,y)(P(x,xyvw) + P(y,xyvw))$. Equation (13) does not assume any choice model, it merely expresses the probability of a constrained choice in terms of the probabilities of non-constrained choices.

A choice model is called invariant if the probability of choice is unaffected by constraints imposed on offered sets. Thus, invariance implies that $P(x,[A][B]) = P(x, A \cup B)$ for all $x \in A \cup B$. It is easy to see that CRM is invariant. In fact, the invariance condition is equivalent to Luce's (1959) choice axiom, which asserts that $P(x,A) = P(x,B)\,P(B,A)$ whenever $B \subset A$ and $P(x,A) > 0$. Consequently, Luce's model is the only invariant theory of choice; all other models violate invariance in one form or another!

Two hierarchical structures or trees defined on the same set of alternatives are called compatible iff there exists a third tree, defined on the same alternatives, which is a refinement of both. Refinement is used here in a non-strict sense so that every tree is a refinement of itself. Thus, ((xy)z)(uvw) is compatible with (xyz)((uv)w) because both are coarsenings of ((xy)z)((uv)w) . On the other hand, (xy)z and (xz)y are incompatible since there is no tree that is a refinement of both. Note that the (degenerate) tree structure implied by CRM is compatible with any tree. The relation between the intrinsic preference tree and the imposed agenda is described in the following theorem.

COMPATIBILITY THEOREM: If (13) holds and Pretree is valid then a set of choice probabilities is unaffected by constraints iff the constraints are compatible with the structure of the tree.

A proof of the theorem is given in Section V of the Appendix; the following discussion explores the simplest example of the effect of agenda. Suppose $T=\{x,y,z\}$, Pretree holds and the intrinsic tree is (xy)z.

Let α, β and γ denote the measures of the unique aspects of x, y and z, respectively, and let θ denote the measure of the aspects shared by x and y, see Figure 5. Setting $\alpha+\beta+\gamma+\theta = 1$, yields

$$P(x,xyz) = \alpha + \theta\alpha / (\alpha+\beta), \; P(y,xyz) = \beta + \theta\beta / (\alpha+\beta), \; P(z,xyz) = \gamma.$$

There are three non-trivial constraints in this case. The first, [xy]z, coincides with the tree structure, hence it does not influence choice probability. The other two partitions, [xz]y and [yz]x, are symmetric with respect to x and y, hence we investigate only the former. By (13), we have P(y, [xz]y) = P(y,xyz). More generally, an imposed partition, e.g., [xz]y, does not change the probability of selecting the isolated alternative, e.g., y. The imposed constraint, however, can have a substantial effect on the probability of selecting other alternatives, e.g., x and z. Since

$$P(x,[xz]y) = P(x,z)\,(P(x,xyz) + P(z,xyz)),$$

$$P(x,[xz]y) > P(x,xyz) \quad \text{iff}$$

$$P(z,xyz)\,P(x,z) > P(x,xyz)\,P(z,x).$$

In the tree model, with (xy)z, this inequality is always satisfied, see Equation (7), because

$$\frac{P(x,z)}{P(z,x)} = \frac{\alpha+\theta}{\gamma} > \frac{\alpha+\theta\alpha/(\alpha+\beta)}{\gamma} = \frac{P(x,xyz)}{P(z,xyz)},$$

hence, P(x,[xz]y) > P(x,xyz). Imposing the partition [xz]y, therefore, on the tree (xy)z is beneficial to x, immaterial for y, and harmful to z.

To interpret this result, recall that x and y share more aspects with each other than with z. In the absence of external constraints, z benefits

directly from the competition between x and y -- as demonstrated by the above inequality which shows that x loses proportionally more than z by the addition of y to the set $\{x,z\}$. The constraint [xz]y reduces, in effect, the direct competition between x and y, and enhances x at the expense of z.

A numerical example illustrates this effect. Suppose $\alpha = .0001$, $\beta = .0999$, $\theta = .4$ and $\gamma = .5$. In a free choice, therefore, P(z, xyz) = .5, P(y,xyz) = .4995 and P(x,xyz) = .0005 because x is practically dominated by y. Under the constraint [xz]y, however, the probabilities of choosing z, y and x, respectively, are .2761, .4995 and .2244. Thus, the imposed partition increases the probability of choosing x from .0005 to .2244 This occurs because x fares well against z, but performs badly against y. In a regular choice where x is compared directly to y, its chances are negligible. Under the partition [xz]y, however, these chances improve greatly because there is an even chance to eliminate y in the first stage, and a close-to-even chance to eliminate z in the second stage.

The above treatment of constrained choice should be viewed as a first approximation because its assumptions probably do not always hold. First, the alternatives in question may not form a tree. Second, the independence condition, embodied in (13), may fail in many situations. Finally, the probability of selecting A over B may not equal $\sum_{x \in A} P(x, A \cup B)$ - particularly when A and B have a different number of elements that could induce a bias to choose the larger or the smaller set. Nevertheless, the proposed model appears to provide a promising method for the analysis of constrained choice.

Constrained Choices among Prospects and Applicants

The present experiment investigates the effect of agenda on individual choice, and tests the implications of the preceding analysis. Two parallel studies are reported using hypothetical prospects (Study I) and college applicants (Study II) as choice alternatives. Each prospect was described as p% chance to win \$a and (100 - p)% chance to win nothing, denoted (\$a,p%). Each applicant was characterized by a high school grade point average (GPA) and an average score on the Scholastic Achievement Test (SAT). The subjects were reminded that the SAT has a maximum of 800 with an average of about 500, and that GPA is computed by letting A = 4, B = 3, etc.

One hundred students from Stanford University participated in each of the two studies. Every subject was presented individually with 10 triples of alternatives, each displayed on a separate card. Each triple was divided into a pair of alternatives and an odd alternative, and the subject was instructed to decide first whether he or she preferred the odd alternative of one of the members of the pair. If the odd alternative was selected, the elements of the triple were not considered again. If the pair was selected, the subject was given an opportunity to choose between its members after the presentation of all ten triples. The delay was designed to reduce the dependence between the trinary and the binary choices.

The subjects in Study I were asked to imagine that they were actually faced with the choice between the displayed prospects, and to indicate

the decision they would have made in each case. The subjects in Study II were asked to select, from each triple, the applicant that they preferred. Subjects were reminded that their task was to express their preferences rather than predict which applicant was most likely to be admitted to college. The participants in both studies were asked to consider each choice carefully and to treat each triple as a separate choice problem.

The alternatives in each triple, denoted x,y,z, were constructed so that (i) x and y are very similar, (ii) z is not very similar to either x or y, (iii) the advantage of y over x on one dimension appears greater than the advantage of x over y on the other dimension, so that y is preferable to x. In Study I, z is a sure prospect while x and y are risky prospects with similar probabilities and outcomes, and with y superior to x in expected value. For example, x = ($40, 75%), y = ($50, 70%) and z is $25 for sure, denoted ($25). In Study II, x and y are applicants with relatively high GPA and moderate SAT, while z is an applicant with a relatively low GPA and fairly high SAT. For example, x = (3.5, 562), y = (3.4, 596) and z = (2.5, 725). The results of a pilot study indicated that one-tenth of a point on the GPA scale is roughly equivalent to twenty SAT points. According to this criterion for overall quality, applicant y is 'better' than x in all cases. All triples of prospects and applicants are displayed in Table 2.

The present experiment was designed to compare choice under [xy]z with choice under [xz]y. Hence, for each triple, one-half of the subjects had to choose first between the pair (x,y) and z, while the remaining one-half had to choose first between the pair (x,z) and y. Each subject made five

choices under [xy]z and five choices under [xz]y. The order of triples and constraints, as well as the positions of the option cards (i.e., left, center, right) were all counterbalanced.

Because alternatives x and y have much more in common with each other than with z, the tree structure that best approximates the triples is (xy)z. Hence, the constraint [xy]z is compatible with the natural structure of the alternatives, while the constraint [xz]y is not. The preceding analysis implies that the latter should enhance the choice of x, hinder the choice of z, and have no substantial effect on the choice of y. Stated formally,

$$d(x) = P(x,[xz]y) - P(x,[xy]z) > 0$$

$$d(y) = P(y,[xz]y) - P(y,[xy]z) = 0$$

$$d(z) = P(z,[xz]y) - P(z,[xy]z) < 0$$

Obviously, in the absence of any effect due to the imposed constraints $d(x) = d(y) = d(z) = 0$. The proportions of subjects that chose x and y in each triple under the two constraints are presented in Table 2, along with the values of d(x), d(y) and d(z) defined above.

Insert Table 2 here

The results reported in Table 2 tend to confirm the predicted pattern of choices. In both studies the values of $d(x)$ are all positive while the values of $d(z)$ are negative with a few small exceptions. Furthermore, in both Studies I and II the means of $d(x)$ are significantly positive, yielding $t(9) = 9.2$ and $t(9) = 8.6$, respectively, $p < .001$, while the means of $d(z)$ are significantly negative, yielding $t(9) = -3.0$, $p < .05$, in Study I, and $t(9) = 5.5$, $p < .001$ in Study II. The means of $d(y)$ were also negative, yielding $t(9) = -2.3$ and $t(9) = -2.8$, respectively, $.01 < p < .05$. Hence, the shift from the natural constraint $[xy]z$ to the constraint $[xz]y$ increases the chances of x and decreases the chances of z and, to a lesser extent, of y. The latter effect, which departs from the predicted pattern, may reflect a response bias against the odd alternative.

The pattern of results described in Table 2 seems to exclude two alternative simple models that produce an agenda effect. Suppose choices are made at random so that one chooses between the odd and the paired alternatives with equal probability. As a consequence,

$$d(x) = P(x,[xz]y) - P(x,[xy]z) = \tfrac{1}{2} \times \tfrac{1}{2} - \tfrac{1}{2} \times \tfrac{1}{2} = 0$$

$$d(y) = P(y,[xz]y) - P(y,[xy]z) = \tfrac{1}{2} - \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{4} > 0, \text{ and}$$

$$d(z) = P(z,[xz]y) - P(z,[xy]z) = \tfrac{1}{2} \times \tfrac{1}{2} - \tfrac{1}{2} = -\tfrac{1}{4} < 0$$

which are incompatible with the experimental findings.

The random choice model gives a distinct advantage to the odd alternative, hence its failure suggests a different model according to which the odd alternative suffers a setback, perhaps because people prefer to delay the choice and avoid commitment. This hypothesis, however, implies $d(x) = 0$, $d(y) < 0$, and $d(z) > 0$ -- again contrary to the data.

Since all triples have the same structure, it is possible to pool all x-choices, y-choices and z-choices across triples and test our hypotheses within the data of each subject. Let $P_i(x,[xz]y)$ denote the proportion of triples in which subject i made an x-choice under the constraint [xz]y, etc. Let $d_i(x) = P_i(x,[xz]y) - P_i(x,[xy]z)$, $d_i(z) = P(z,[xz]y) - P_i(z,[xy]z)$, and let $D_i = d_i(x) - d_i(z)$. Thus, D_i measures the advantage of x over z due to the shift from [xy]z to [xz]y. Recall that, in the absence of an agenda effect $d_i(x) = d_i(z) = D_i = 0$, while under the proposed model $d_i(x) > 0 > d_i(z)$ and hence D_i has a positive expectation. The means of the D_i distributions are .21 in Study I and .25 in Study II, which are significantly positive, yielding $t(99) = 4.2$, and $t(99) = 5.8$, respectively, $p < .001$ in both cases. In Study I, 60% of the D_is are positive and 22% negative; in the Study II, 62% are positive and 18% negative. Hence, the predicted pattern of choices is also confirmed in a within-subject comparison, where choices are pooled over trials rather than over subjects.

In summary, the data show that imposed constraints have a significant impact on choice behavior, and that the results confirm the major predictions of the proposed model of constrained choice. The present results about individual choice, that are based on the correlational pattern among the alternatives, should be distinguished from the results of Plott and Levine (1978) who demonstrated the effect of agenda on the outcome of group decision based on majority vote. An agenda often introduces strategic considerations that could affect the outcome of a voting process, even if it does not change the ordering of the options for any single individual, much as group decision can be intransitive even when its members are all transitive. Although different effects seem to

contribute to the failure of invariance in individual and in collective choice, they are probably both present, for example, in many forms of committee decision making. The influence of procedural constraints on either individual or social choice emerges as a subject of great theoretical and practical significance. For if the choice of a new staff member, for example, depends on whether the initial decision concerns the nature of the appointment (e.g., junior vs. senior), or the field (e.g., perception vs social), then the order in which decisions are made becomes an important component of the choice process that cannot be treated merely as a procedural matter.

The present model of individual choice under constraints may serve three related functions. First, it could be used to predict the manner in which choices among political candidates, market products or public policies are affected by the introduction or the change of agendas. Second, the model may be used to construct an agenda so as to maximize the probability of a desired outcome. Experienced politicians and seasoned marketeers are undoubtedly aware of the effects of grouping and separating options. A formal model may nevertheless prove useful, particularly in complex decisions where the number of alternatives is large and computational demands exceed cognitive limitations. Third, the model can be employed by a group or a committee as a framework for the discussion and comparison of different agendas. Although an 'optimal' or a 'fair' agenda may not exist, the analysis might help clarify the issues and facilitate the choice. If all members of the group, for example, perceive the available options in terms of the same tree structure, even though they have different weights and preferences, then the use of an agenda that is compatible with that structure is recommendable since it ensures invariance. The applications of the present development for the construction, selection, and evaluation of agendas are still left to be developed.

DISCUSSION

Individual choice behavior is variable, complex and context dependent, and the attempts to model it are, at best, incomplete. Even the most basic axioms of preference are consistently violated under certain circumstances, see, e.g., Kahneman and Tversky (1979), Lichtenstein and Slovic (1968), Tversky (1969). The present treatment does not attempt to develop a comprehensive theory of choice, but rather to analyze in detail a particular strategy that appears to govern several decision processes. There are undoubtly decision processes that are not compatible with Pretree. Some of them could perhaps be explained by EBA, while others may require different theoretical treatments. The selection of a choice model, however, generally involves a balance between generality or scope on the one hand, and simplicity or predictive power on the other. Pretree may be regarded as an intermediate model that is much less restrictive than CRM since it is compatible with the similarity hypothesis, yet it is much more parsimonious than the general EBA model since it has at most $2n - 2$ rather than $2^n - 2$ parameters.

Furthermore, the tree model may provide a useful approximation to a more complex structure, in the same way that a two dimensional solution often provides a useful representation of a higher dimensional structure. Consider, for example, a person who is about to take a one-week trip to a single European country and is offered a choice between France (F) and Italy(I) and between a luxury tour(L) and an economy tour(E). Naturally, the luxury tour is much more comfortable but also considerably more expensive than the economy tour. It is easy to see that the four available alternatives F_L, F_E, I_L, I_E do not satisfy the inclusion rule because, for any triple, each alternative shares different aspects with the other two. Hence, the EBA model cannot be reduced to a tree in this case, although it can be approximated by a tree -- provided one of the

attributes looms much larger than the others.

Suppose the decision maker is very concerned about the site of the trip (Italy vs. France) but is not overly concerned about comfort or price. In this case, the weights associated with the tour-type (luxury vs. economy) would be small in comparison with the weights associated with the sites. Hence, the observed choice probabilities could be approximated fairly well by the tree $(F_L F_E)$ $(I_L I_E)$. On the other hand, if the decision maker is much more concerned about the type of tour than about its site, his choice probabilities will be better described by the tree $(F_L I_L)$ $(F_E I_E)$. The quality of either approximation depends on the degree to which one attribute dominates the other, and it could be assessed directly by examining the trinary and the quarternary conditions. An extension of the tree model that deals with factorial structures will be described elsewhere.

Hierarchical or tree-like models of choice have been recently employed by students of economics and market research who investigate questions such as the share of the market to be captured by a new product, or the probability that a consumer will switch from one brand to another. Luce's model provides the simplest answers to such questions, but as we have already seen, it is too restrictive. Perhaps the simplest way of extending CRM is to assume that the offered set of alternatives can be partitioned into classes so that the model holds within each homogeneous class, even though it does not hold for heterogeneous sets.

This assumption underlies the analysis of brand switching developed by the Hendry Corporation, and described by Kalwani and Morrison (1977). According to the Hendry model, the probability that a consumer will purchase a new brand given that he switched from his old one, is proportional to the market share of the new brand-- provided the two brands

belong to the same class of the partition. The application of this model, therefore, requires prior identification of an appropriate partition, or tree structure, that is presumably constructed on the basis of informed intuition. The similarity-based scaling procedure employed in this paper, and the test of the necessary trinary and quarternary conditions could perhaps be used to construct and validate the partition to which the analysis of brand switching is applied.

The partition of the alternatives into homogeneous classes satisfying CRM was also used by McFadden (1976, 1978) in his theoretical and empirical analyses of probabilistic choice. As an economist, McFadden was primarily interested in aggregate demand for alternatives (e.g., different modes of transportation) as a function of measured attributes of the alternatives and the decision makers (e.g., cost, travel time, income). The Thurstonian, or random utility, model provides a natural framework for such an analysis which assumes, in accord with classical economic theory, that each individual maximizes his utility function defined over the relevant set of alternatives and the random component reflects the sampling of individuals with different utility functions.

McFadden (1978) began with the multinomial logit (MNL) model in which

$$P(x,A) = \exp \sum_i x_i \theta_i \Big/ \sum_{y \in A} \exp \sum_i y_i \theta_i$$

where $x_1, \ldots, x_n$ are specified attributes of x, and $\theta_1, \ldots, \theta_n$ are parameters estimated from the data. This is clearly a special case of Luce's model(2), where log u(x) is a linear function in the parameters $\theta_1, \ldots, \theta_n$. It is expressible as a random utility model by assuming an extreme value distribution $F(t) = \exp[-\exp-(at+b)], a > 0$, see e.g., Luce (1977), Yellott(1977).

The MNL model has been applied to several economic problems, notably transportation planning (McFadden, 1976), but the failure of context-independence led McFadden (1978) to develop a more general family of choice

models, called generalized extreme value models, that are compatible with the similarity hypothesis. One model from this family, called the nested logit model, assumes a tree structure in which the probabilities of choice at each level of the tree conform to the multinomial logit model, see McFadden (1978). Although the nested logit model does not coincide with Pretree, the two models are sufficiently close that the former may be regarded as a random utility counterpart of the latter.

Psychological models of individual choice fall into three overlapping classes: decomposition models, probabilistic models and process models. Decomposition models express the overall value of each alternative as a function of the scale values associated with its components. This class includes all the variations of expected utility theory as well as the various adding and averaging models. Probabilistic models relate choice data to an underlying value structure through a probabilistic process. The models of Thurstone and Luce are prominent examples. Process models attempt to capture the mental operations that are performed in the course of a decision. This approach, pioneered by Simon, has led to the development of computer models designed to simulate the decision making process. Pretree, like the more general EBA, belongs to all three classes. It is a decomposition model that expresses the overall value of an alternative as an additive combination of the values of its aspects. Unlike most decomposition models, however, the relation between the observed choice and the underlying value structure is probabilistic , and the formal theory is interpretable as a process model of choice behavior that is based on successive eliminations following a tree structure.

This paper exhibits three correspondence relations (i) the equivalence of elimination-by- tree and the hierarchical elimination model, (ii) the compat-

ibility of aggregate choice and the individual EBA model, and (iii) the correspondence between preference and similarity trees. The three results, however, have different theoretical and empirical status. The equivalence of EBT and HEM is a mathematical fact that permits the application of the tree model to both random and hierarchical decision processes. The second result offers a new interpretation of EBA as an aggregate choice model, thereby providing a rationale for applying EBA to aggregate data. Finally, the compatibility of similarity and preference trees is an empirical observation which suggests that the two processes are related through a common underlying structure.

MATHEMATICAL APPENDIX

I. Proof ot the Structure Theorem

To show that a tree representation of $T^* = \{x' | x \varepsilon T\}$ implies the inclusion rule, let $t(x)$ denote the path from the root of the tree to the terminal node associated with x. For any x, y, z, in T there are 4 possible tree structures, and they all satisfy the inclusion rule as shown below.

a. If $t(x)$ and $t(y)$ meet below $t(z)$, then $x' \cap y' \supset x' \cap z'$.

b. If $t(x)$ and $t(z)$ meet below $t(y)$, then $x' \cap z' \supset x' \cap y'$.

c. If $t(y)$ and $t(z)$ meet below $t(x)$, then $x' \cap y' = x' \cap z'$.

d. If $t(x)$, $t(y)$ and $t(z)$ all meet at the same node then $x' \cap y' = x' \cap z'$.

In order to establish the sufficiency of the inclusion rule, let $T_\alpha = \{x \varepsilon T | \alpha \varepsilon x'\}$, and let $S(T)$ be the set of all T_α for any α in T'. To prove that $T^* = \{x' | x \varepsilon T\}$ is a tree, it suffices to show that $S(T)$ is a hierarchical clustering. That is, for any α, β in T' either $T_\alpha \supset T_\beta$, or $T_\beta \supset T_\alpha$, or $T_\alpha \cap T_\beta$ is empty. Suppose $S(T)$ is not a hierarchical clustering. Then there exist some distinct aspects α, β in T' and some x, y, z in T such that $x \varepsilon T_\alpha \cap T_\beta$, $y \varepsilon T_\alpha - T_\beta$ and $z \varepsilon T_\beta - T_\alpha$. Hence, α is included in $x' \cap y'$, β is included in $x' \cap z'$, but α is not included in z' and β is not included in y'. Consequently, $x' \cap y'$ neither includes nor is included in $x' \cap z'$ and the inclusion rule is violated, as required.

II. Proof of the Equivalence Theorem.

(i) EBT implies HEM.

If EBT holds for T, then it must also hold for any $A \subset T$ with the induced tree structure. Hence, it suffices to demonstrate the first two parts of Equation (3)

(a) If $\gamma|\beta$ and $\beta|\alpha$ then $P(A_\alpha, A_\gamma) = P(A_\alpha, A_\beta)P(A_\beta, A_\gamma)$.

(b) If $\gamma|\beta$ and $\gamma|\alpha$ then $\dfrac{P(A_\alpha, A_\gamma)}{P(A_\beta, A_\gamma)} = \dfrac{m(\alpha)}{m(\beta)}$, provided $m(\beta) \neq 0$.

We begin with the following auxiliary result. If $\beta|\alpha$, then

$$P(x, A_\beta) = P(x, A_\alpha)\frac{m(\alpha)}{m(\beta)-u(\beta)}$$

Let $\alpha_1, \ldots, \alpha_n$ be a sequence of links leading from x to α. That is, $A_{\alpha_1} = \{x\}$, $\alpha_{i+1}|\alpha_i$, $i=1,\ldots,n-1$, and $\alpha_n = \alpha$. Assuming EBT and $\beta|\alpha$

$$P(x, A_\beta) = \frac{u(\alpha_n)}{m(\beta)-u(\beta)}P(x, A_{\alpha_n}) + \frac{u(\alpha_{n-1})}{m(\beta)-u(\beta)}P(x, A_{\alpha_{n-1}}) + \ldots + \frac{u(\alpha_1)}{m(\beta)-u(\beta)}P(x, A_{\alpha_1})$$

$$= \frac{u(\alpha_n)}{m(\beta)-u(\beta)}P(x, A_{\alpha_n}) + \frac{m(\alpha_n)-u(\alpha_n)}{m(\beta)-u(\beta)}\left(\frac{u(\alpha_{n-1})}{m(\alpha_n)-u(\alpha_n)}P(x, A_{\alpha_{n-1}}) + \ldots + \frac{u(\alpha_1)}{m(\alpha_n)-u(\alpha_n)}P(x, A_{\alpha_1})\right)$$

$$= \frac{u(\alpha_n)}{m(\beta)-u(\beta)}P(x, A_{\alpha_n}) + \frac{m(\alpha_n)-u(\alpha_n)}{m(\beta)-u(\beta)}P(x, A_{\alpha_n})$$

$$= \frac{m(\alpha)}{m(\beta)-u(\beta)}P(x, A_\alpha).$$

as required. To prove (b) we assume that $\gamma|\beta$ and $\gamma|\alpha$, hence

$$\frac{P(A_\alpha,A_\gamma)}{P(A_\beta,A_\gamma)} = \frac{\sum\limits_{x\varepsilon A_\alpha} P(x,A_\gamma)}{\sum\limits_{x\varepsilon A_\beta} P(x,A_\gamma)}$$

$$= \frac{\sum\limits_{x\varepsilon A_\alpha} P(x,A_\alpha)\dfrac{m(\alpha)}{m(\gamma)-u(\gamma)}}{\sum\limits_{x\varepsilon A_\beta} P(x,A_\beta)\dfrac{m(\beta)}{m(\gamma)-u(\gamma)}}$$

$$= \frac{m(\alpha)}{m(\beta)}$$

since $\sum\limits_{x\varepsilon A_\alpha} P(x,A_\alpha) = \sum\limits_{x\varepsilon A_\beta} P(x,A_\beta) = 1.$

To prove (a), suppose $\gamma|\beta$ and $\beta|\alpha$. By our auxiliary result $P(A_\beta,A_\gamma) = \sum\limits_{x\varepsilon A_\beta} P(x,A_\gamma) = m(\beta)/(m(\gamma)-u(\gamma))$, and

$$P(A_\alpha,A_\gamma) = \sum_{x\varepsilon A_\alpha} P(x,A_\gamma)$$

$$= \sum_{x\varepsilon A_\alpha} P(x,A_\beta)\,\frac{m(\beta)}{m(\gamma)-u(\gamma)}$$

$$= P(A_\alpha,A_\beta)\,\frac{m(\beta)}{m(\gamma)-u(\gamma)}$$

$$= P(A_\alpha,A_\beta)\,P(A_\beta,A_\gamma)$$

(ii) HEM implies EBT.

We have to show that for any $A \subset T$, $P(x,A)$ satisfies Equation (1). The proof is by induction on the cardinality of A. Let $\alpha_1,\ldots,\alpha_n$ be the sequence of segments leading from x to the root of A. That is, $\{x\} = A_{\alpha_1}$, $\alpha_{i+1}|\alpha_i$, $i = 1,\ldots, n-1$, and $A_{\alpha_n} = A$. If $\gamma|\beta$, $x\varepsilon A_\beta$, and Equation (3) holds then

$$P(x,A_\gamma) = \frac{m(\beta)}{m(\gamma)-u(\gamma)}\,P(x,A_\beta)$$

Thus, using the inductive hypothesis, we obtain

$$P(x,A_{\alpha_n}) = \frac{m(\alpha_{n-1})}{m(\alpha_n)-u(\alpha_n)} P(x,A_{\alpha_{n-1}})$$

$$= \frac{u(\alpha_{n-1})}{m(\alpha_n)-u(\alpha_n)} P(x,A_{\alpha_{n-1}}) + \frac{m(\alpha_{n-1})-u(\alpha_{n-1})}{m(\alpha_n) - u(\alpha_n)} P(x,A_{\alpha_{n-1}})$$

$$= \frac{u(\alpha_{n-1})}{m(\alpha_n)-u(\alpha_n)} P(x,A_{\alpha_{n-1}}) + \frac{m(\alpha_{n-1})-u(\alpha_{n-1})}{m(\alpha_n) - u(\alpha_n)} \cdot \frac{\sum_{i=1}^{n-2} u(\alpha_i)P(x,A_{\alpha_i})}{m(\alpha_{n-1})-u(\alpha_{n-1})}$$

$$= \frac{\sum_{i=1}^{n-1} u(\alpha_i)\, P(x,A_{\alpha_i})}{m(\alpha_n) - u(\alpha_n)}$$

which is the recursive expression for $P(x,A~)$.

III. Proof of the Representation Theorem.

The proof is divided into a series of lemmas. Let P_T denote the set of binary choice probabilities defined for all pairs of elements in T.

Lemma 1: If $T = \{x,y,z\}$, then P_T satisfies Pretree with (xy)z iff the trinary inequality (4) is satisfied in this form.

Proof: Necessity is obvious. To prove sufficiency, we use the notation of Figure 5, where $R(x,y) \geq 1$. Set $\alpha = 1$, $\beta = R(y,x)$, and select $\theta \geq 0$ so that $[R(x,z) - R(y,z)]\theta = R(y,z) - R(y,x)\ R(x,z)$, and let $\gamma = R(z,x)(1+\theta)$. (Note that when $R(x,y) > 1$, θ is uniquely defined and positive, and when $R(x,y) = 1$, θ can be chosen arbitrarily).

Let $\bar{P}_T$ be the set of binary probabilities obtained by using the above expressions for $\alpha,\beta,\gamma,\theta$ in the defining equations of the model. It can be verified, after some algebra, that $\bar{P}_T = P_T$ as required.

Before we go further, note that if P_T satisfies Pretree with (xy)z and $R(x,y) > 1$ then $\beta/\alpha = R(y,x)$. Furthermore,

$$\frac{\frac{\theta}{\alpha} + 1}{\frac{\theta}{\alpha} + \frac{\beta}{\alpha}} = \frac{\theta+\alpha}{\theta+\beta} = \frac{R(x,z)}{R(y,z)} \text{ implies } \frac{\theta}{\alpha} = \frac{R(y,z) - R(y,x)R(x,z)}{R(x,z) - R(y,z)} \text{ and}$$

$$\frac{1+\frac{\theta}{\alpha}}{\frac{\gamma}{\alpha}} = \frac{\alpha+\theta}{\gamma} = R(x,z) \text{ implies } \frac{\gamma}{\alpha} = R(z,x)\left(1+\frac{R(y,z) - R(y,x)\ R(x,z)}{R(x,z) - R(y,z)}\right) = \frac{1-R(y,x)}{R(x,z)-R(y,z)}.$$

Hence, the lengths of all the links are determined up to multiplication by a positive constant. Furthermore, the present model readily entails the following property.

Lemma 2: Suppose A and B = {x,y,v} are sets of objects such that $y,v \in A$ and $x \notin A$, and suppose that both P_A and P_B satisfy Pretree. (It is assumed that P(v,y) is the same in both structures). Then the measures on A' and B' can be selected so that u(v'-y') -- as well as u(y'-v') -- are the same in both measures.

Lemma 3: Suppose A = {x,y,v} and B = {y,v,w} satisfy Pretree, with representing measures u_A and u_B, in the forms (xy)v and (yv)w, respectively. If $C = A \cup B$ = {x,y,v,w} satisfies the appropriate quarternary condition with (xy)(v,w) or with ((xy)v)w, then there exists a representing measure u on C' which extends both u_A and u_B. Naturally, we assume that u_A and u_B were selected according to Lemma 2.

Proof: Consider the form (xy)(vw), see Figure 5a. By Lemma 2, $u_A(\beta+\theta) = u_B(\beta+\theta)$ and $u_A(\lambda+\gamma) = u_B(\lambda+\gamma)$. Hence, u_A and u_B can be used to define a measure u on C'. To show that u is a representing measure on C' we have to show that $R(x,w) = u(\theta+\alpha)/u(\lambda+\delta)$. Since C satisfies Pretree, it follows from (5) that

$$R(x,w) = R(y,w)R(x,v)R(v,y)$$

$$= \frac{u(\beta+\theta)}{u(\lambda+\delta)}\frac{u(\alpha+\theta)}{u(\lambda+\gamma)}\frac{u(\lambda+\gamma)}{u(\beta+\theta)}$$

$$= \frac{u(\theta+\alpha)}{u(\lambda+\delta)}$$

Next, consider the form ((xy)v)w, see Figure 5b. Here, we have to show that $R(x,w) = u(\alpha+\theta+\lambda)/u(\delta)$. Applying (6) it follows that

$$R(x,w) = \frac{(1-R(x,v))R(y,w) + R(v,w)(R(x,v) - R(y,v))}{1-R(y,v)}$$

$$= \frac{\left(1-\frac{u(\alpha+\theta)}{u(\gamma)}\right)\frac{u(\beta+\theta+\lambda)}{u(\delta)} + \frac{u(\gamma+\lambda)}{u(\delta)}\left(\frac{u(\alpha+\theta)}{u(\gamma)} - \frac{u(\beta+\theta)}{u(\gamma)}\right)}{1-\frac{u(\beta+\theta)}{u(\gamma)}}$$

$$= \frac{u(\alpha+\theta+\lambda)}{u(\delta)} \quad \text{as required.}$$

Lemma 4: P_T satisfies Pretree with a specified structure iff for every $S \subset T$, with four elements or less, P_S satisfies Pretree relative to the same structure.

Proof: Necessity is immediate. Sufficiency is proved by induction on the cardinality of T, denoted n. Suppose n > 4, and assume that the lemma holds for any cardinality smaller than n.

Suppose (xy)v holds for any v in T. Let $A = T - \{x\}$, and $B = \{x,y,v\}$. By the induction hypothesis, both P_A and P_B satisfy Pretree with the appropriate structure. By Lemma 2 we can assume, with no loss of generality, that the measures of y and v in A' coincide with their measures in B'. Since any aspect in T' appears either in A' or in B', and since the aspects that appear in both trees have the same measure, we can define the measure of any aspect in T' by its measure in A' or in B'. Letting $\bar{P}$ denote the calculated binary probability function, we show that $\bar{P}_T = P_T$.

Since $\bar{P}_A = P_A$ and $\bar{P}_B = P_B$, it remains to be shown that $\bar{P}(x,w) = P(x,w)$ for any $w \in T-B$.

Let $C = \{x,y,v,w\}$, which satisfies Pretree, by assumption, with either (xy)(vw) or ((xy)v)w. Since $C = B \cup \{y,v,w\}$, Lemma 3 implies that the representing measure

on C' coincides with the restriction to C' of the defined measure on T'. Hence, $\bar{P}(x,w) = P(x,w)$ as required.

In conclusion, Lemma 3 together with Lemma 1 show that the trinary and the quarternary conditions are necessary and sufficient for the representation of quadruples. Lemma 4 shows that if Pretree is satisfied by all quadruples, then it is satisfied by the entire object set. This completes the proof of the representation theorem.

IV. Uniqueness Considerations.

It follows readily from the representation theorem that, given a tree structure, the measure u is unique up to multiplication by a positive constant except in the case where ~~all~~ binary choice probabilities equal 1/2. We show that the tree structure is uniquely determined by the binary and the trinary choice probabilities, but not by the binary data alone.

To show that binary choice probabilities do not always determine a unique tree structure, consider two different trees (xy)z and (yz)x, and let α, β, γ denote, respectively, the unique aspects of x, y, z, let θ denote the aspects shared by x and y, and let λ denote the aspects shared by y and z. Let u and v be the measures associated with (xy)z and (yz)x, respectively, and suppose that

$$u(\alpha) = 2,\ u(\beta) = 1,\ u(\gamma) = 1, \text{ and } u(\theta) = 2$$

$$v(\alpha) = 8,\ v(\beta) = 3,\ v(\gamma) = 1, \text{ and } v(\lambda) = 1$$

By the assumed tree structures $u(\lambda) = v(\theta) = 0$. It is easy to verify that the two trees yield identical binary choice probabilities: $P(x,y) = 2/3$, $P(y,z) = 3/4$, $P(x,z) = 4/5$. We next show that the tree structure is uniquely determined by the binary <u>and</u> the trinary choice probabilities, provided all binary probabilities are non-zero. Consider a tree (xy) z with a measure u, and aspects $\alpha, \beta, \gamma, \theta$

defined as above. Assume $u(\alpha)$, $u(\beta)$, $u(\gamma)$ and $u(\theta)$ are nonzero. It follows from (xy)z that

$$\frac{P(x,y)}{P(y,x)} = \frac{u(\alpha)}{u(\beta)} = \frac{u(\alpha) + u(\theta)u(\alpha)/(u(\alpha) + u(\beta))}{u(\beta) + u(\theta)u(\beta)/(u(\alpha) + u(\beta))} = \frac{P(x,xyz)}{P(y,xyz)}$$

Suppose the data were compatible with another tree structure, say (yz)x with no loss of generality. By the same argument

$$\frac{P(y,z)}{P(z,y)} = \frac{P(y,xyz)}{P(z,xyz)}, \text{ and hence}$$

$$\frac{u(\beta) + u(\theta)}{u(\gamma)} = \frac{u(\beta) + u(\theta)u(\beta)/(u(\alpha) + u(\beta))}{u(\gamma)}$$

which implies $u(\alpha) = 0$ contrary to our assumption. Given both binary and trinary probabilities, therefore, the structure of any triple and hence of the entire tree is uniquely determined.

V. Proof of the Compatability Theorem.

It follows readily from HEM, see Equation (3), that

$$P(x,A) = P(x,A_1)P(A_1,A_2)\ldots P(A_{n-1},A_n)$$

for some sequence $A_1,\ldots,A_n$ such that $A_n=A$, and $A_i \subset A_{i+1}$, $i=1,\ldots,n-1$. We show first that the sequence can be chosen so that $a_i=i+1$, $1 \leq i \leq n$, where a_i is the cardinality of A_i. This condition is obviously satisfied in a binary tree where each node joins at most two links. Suppose then that the tree contains three links that meet at the same node, e.g., $\delta|\gamma$, $\delta|\beta$ and $\delta|\alpha$. Hence, by part (b) of Equation (3),

$$P(A_\alpha,A_\delta) = \frac{m(\alpha)}{m(\alpha)+m(\beta)+m(\gamma)} = \frac{m(\alpha)}{m(\alpha)+m(\beta)} \times \frac{m(\alpha)+m(\beta)}{m(\alpha)+m(\beta)+m(\gamma)} =$$

$$= P(A_\alpha,A_\alpha \cup A_\beta)P(A_\alpha \cup A_\beta,A_\delta),$$

and the result is readily extended to nodes with k links. Under Pretree, therefore, $P(x,A)$ is expressible as a product where each factor $P(A_i,A_{i+1})$ is a probability of choosing between two branches.

Under Equation (13), the probability of selecting x from A under a specified agenda equals $P(x,B_1)P(B_1,B_2)\ldots P(B_m,A)$, for some $B_1 \subset B_2 \ldots \subset B_m \subset A$. By compatibility, there exists a tree and hence a binary tree that refines both the adenda and the intrinsic tree structure. By the above argument, $P(x,A)$ is expressible as a product $P(x,A_1)P(A_1,A_2)\ldots P(A_{n-1},A_n)$ where $a_i=i+1$, $1 \leq i \leq n$, corresponding to a binary tree that refines both structures. Thus, each B_j, $j=1,\ldots,m$, appears among the A_i's, $i=1,\ldots,n$. Suppose $B_j = A_i$ and $B_{j+1} = A_{i+t}$, hence

$$P(B_j,B_{j+1}) = P(A_i,A_{i+t}) = \prod_{k=i}^{i+t-1} P(A_k,A_{k+1}), \text{ and}$$

$$P(x,A) = P(x,A_1)P(A_1,A_2)\ldots P(A_{n-1},A_n) = P(x,B_1)P(B_1,B_2)\ldots P(B_m,A).$$

Hence, choice probability is unaffected by an agenda that is compatible with the intrinsic structure of a preference tree.

If the agenda is not compatible with the intrinsic tree, there exists some x,y,z in T such that both (xy)z and [xz]y hold. It is easy to verify (see the discussion in the text) that $P(x,xyz) \neq P(x,[xz]y)$ in this case, which establishes the necessity of the compatability condition.

Reference Notes

1. Falmagne, J. C. Probabilistic theories of measurement. Paper presented at the Soviet-American Seminar on Decision Models, Tbilisi, Soviet Union, April 1979.

References

Carroll, J. D. Individual Differences and Multidimensional Scaling. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences, I. New York: Seminar Press, 1972.

Chandler, J. P. STEPIT--Finds local minima of a smooth function of several parameters. Behavioral Science, 1969, 14, 81-82.

Coombs, C. H. On the use of inconsistency of preferences in psychological measurement. Journal of Experimental Psychology, 1958, 55, 1-7.

Coombs, C. H. A theory of data. New York: Wiley, 1964.

Debreu, G. Review of R. D. Luce, Individual choice behavior: A theoretical analysis. American Economic Review, 1960, 50, 186-188.

Edgell, S. E. Geisler, W. S. III, & Zinnes, J. L. A note on a paper by Rumelhart and Greeno. Journal of Mathematical Psychology, 1973, 10, 86-90.

Kahneman, D., & Tversky, A. Prospect Theory: An analysis of decision under risk. Econometrica, 1979, 47, 263-291.

Kalwani, M. V., & Morrison, D. G. A parsimonious description of the Hendry system. Management Science, 1977, 23, 467-477.

Levine, M. E., & Plott, C. R. Agenda influence and its implications. Virginia Law Review, 1977, 63, 561-604.

Lichtenstein, S., & Slovic, P. Reversal of preference between bids and choices in gambling decisions. Journal of Experimental Psychology, 1971, 89, 46-55.

Luce, R. D. Individual choice behavior: A theoretical analysis. New York: Wiley, 1959.

Luce, R. D. The choice axiom after twenty years. Journal of Mathematical Psychology, 1977, 15, 215-233.

Luce, R. D., & Suppes, P. Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), Handbook of mathematical psychology, III. New York: Wiley, 1965.

Marschak, J. Binary-choice constraints and random utility indicators. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), Mathematical methods in the social sciences, Stanford: Stanford University Press, 1960.

McFadden, D. Quantal choice analysis. A survey. Annals of Economic and Social Measurement, 1976, 5, 363-390.

McFadden, D. Econometric models for probabilistic choice. Cowles Foundation Discussion paper, Yale University, 1978.

Nygren, T. E., & Jones, L. E. Individual differences in perception and preferences for political candidates. Journal of Experimental Social Psychology, 13, 182-197.

Plott, C. R., & Levine, M. E. A model of agenda influence on committee decisions. American Economic Review, 1978, 68, 146-160.

Restle, F. Psychology of judgment and choice. New York: Wiley, 1961.

Rumelhart, D. L., & Greeno, J. G. Similarity between stimuli: An experimental test of the Luce and Restle choice models. Journal of Mathematical Psychology, 1971, 8, 370-381.

Sattath, S., & Tversky, A. Unite and conquer: A multiplicative inequality for choice probabilities. Econometrica, 1976, 44, 79-89.

Sattath, S., & Tversky, A. Additive similarity trees. Psychometrika, 42, 1977, 319-345.

Simon, H. A. Models of man. New York: Wiley, 1957.

Sjöberg, L., & Capozza, D. Preference and cognitive structure of Italian political parties. Italian Journal of Psychology, 1975, 2, 391-402.

Sjöberg, L. Choice frequency and similarity. Scandinavian Journal of Psychology, 1977, 18, 103-115.

Stefflre, V. Some applications of multidimensional scaling to social science problems. In Romney A. K., Shepard R. N., & S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences, II. New York: Seminar Press, 1972.

Thurstone, L. L. A law of comparative judgment. Psychological Review, 1927 34, 273-286.

Tversky, A. Intransitivity of Preferences. Psychological Review, 1969, 76, 31-48.

Tversky, A. Elimination by aspects: a theory of choice. Psychological Review, 1972, 79, 281-299. (a)

Tversky, A. Choice by elimination. Journal of Mathematical Psychology, 1972, 9, 341-367. (b)

Yellott, J. I., Jr. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. Journal of Mathematical Psychology, 1977, 15, 109-144.

Footnotes

[1]The present notion of a preference tree should be distinguished from the concept of a decision tree, commonly used in the analysis of decisions under uncertainty.

[2]*To obtain compact figures* we use a *heavy* line (see Figure 7) to indicate double lenght, and an extra heavy line (see Figure 11) to indicate ten-fold length.

Acknowledgments

This paper has benefited from discussions with R. M. Dawes, J. C. Falmagne, W. Hutchinson, D. H. Krantz, D. Kahneman and D. McFadden. We are particularly grateful to L. Sjöberg for making his data available to us, and to I. Gati and P. Smith for their help in the analysis of the data. This work was supported by the Office of Naval Research under contract N00014-79-0077 to Stanford University.

Table 1

Summary Statistics for the Comparison of CRM and Pretree

Study / Alternatives	Trinary-Inequality Statistics					CRM			Pretree			Difference	
	N	π	R	R_1	R_3	d	χ^2	df	d	χ^2	df	χ^2	df
Rumelhart & Greeno (1971)													
Celebrities	54	89%	1.40	1.13	1.68	.036	78.2^*	28	.023	30.0	25	48.2^*	3
Sjöberg (1977)													
Swedish Parties	23	96%	1.73	1.38	2.27	.038	49.1^*	15	.012	5.8	11	43.3^*	4
Sjöberg & Cappoza (1975)													
Italian Parties	18	78%	1.74	0.93	2.78	.042	67.6^*	15	.023	19.5	11	48.1^*	4
Sjöberg (1977)													
Social Sciences	86	84%	1.52	1.21	1.86	.035	69.1	55	.025	45.5	50	23.6^*	5
Coombs (1958)													
Shades of Gray													
Subject 1	99	83%	2.06	1	4								
Subject 2	139	76%	5.08	1	9								
Subject 3	127	70%	1.45	0.58	3.52								
Subject 4	184	94%	6.66	2.84	100								

† denotes $p < .05$

* denotes $p < .001$

Table 2

Probabilities of Choice Among Prospects and Applicants Under Two Different Constraints

ALTERNATIVES				CONSTRAINTS				EFFECTS		
I Prospects				[xy]z		[xz]y				
Triple	x ($,%)	y ($,%)	z ($)	P(x,[xy]z)	P(z,[xy]z)	P(x,[xz]y)	P(z,[xz]y)	d(x)	d(y)	d(z)
1	(40,75)	(50,70)	(25)	.08	.22	.18	.20	.10	-.08	-.02
2	(80,15)	(75,20)	(10)	.12	.40	.24	.32	.12	-.04	-.08
3	(65,90)	(75,85)	(55)	.12	.42	.20	.46	.08	-.12	.04
4	(120,5)	(85,10)	(5)	.08	.54	.18	.38	.10	.06	-.16
5	(75,30)	(100,25)	(20)	.04	.54	.20	.48	.16	-.10	-.06
6	(125,35)	(120,40)	(35)	.04	.44	.18	.32	.14	-.02	-.12
7	(30,65)	(40,60)	(15)	.18	.36	.30	.40	.12	-.16	.04
8	(35,95)	(45,90)	(30)	.06	.40	.16	.28	.10	.02	-.12
9	(50,85)	(60,80)	(40)	.04	.48	.22	.30	.18	.00	-.18
10	(65,25)	(95,20)	(15)	.02	.42	.24	.28	.22	-.08	-.14
Mean				.078	.422	.210	.342	.132	-.052	-.080
II Applicants										
Triple	x (GPA,SAT)	y (GPA,SAT)	z (GPA,SAT)							
1	(3.3,654)	(3.2,692)	(2.2,773)	.16	.40	.30	.32	.14	-.06	-.08
2	(3.6,592)	(3.5,625)	(2.6,785)	.18	.38	.28	.26	.10	.02	-.12
3	(3.5,579)	(3.7,571)	(2.5,701)	.00	.48	.18	.40	.18	-.10	-.08
4	(3.1,602)	(3.0,641)	(2.1,730)	.14	.36	.22	.26	.08	.02	-.10
5	(2.9,521)	(3.1,515)	(2.3,703)	.04	.50	.20	.34	.16	.00	-.16
6	(2.8,666)	(2.9,661)	(2.0,732)	.06	.40	.26	.24	.20	-.04	-.16
7	(3.8,587)	(3.7,629)	(2.6,744)	.14	.38	.28	.30	.14	-.06	-.08
8	(3.4,600)	(3.6,590)	(2.4,755)	.06	.40	.24	.30	.18	-.08	-.10
9	(3.7,718)	(3.9,712)	(3.1,798)	.00	.40	.20	.22	.20	-.02	-.18
10	(3.5,562)	(3.4,596)	(2.5,725)	.20	.28	.26	.30	.06	-.08	.02
Mean				.098	.398	.242	.294	.144	-.040	-.104

Figure Captions

Figure 1. Schematic representation of three alternatives.

Figure 2. Tree representation of the choice among entrees.

Figure 3. An illustration of the inclusion rule $x' \cap y' \supset x' \cap z'$ (a) as a Venn-diagram, (b) as a tree.

Figure 4. A preference tree for the choice among modes of transportation.

Figure 5. A preference tree for three alternatives.

Figure 6. Preference trees for four alternatives.

Figure 7. Preference tree for choice among celebrities.

Figure 8. Additive tree (ADDTREE) representation of the similarities between Swedish political parties.

Figure 9. Preference tree for choice among Swedish political parties.

Figure 10. Preference tree for choice among Italian political parties.

Figure 11. Preference tree for choice among social sciences.

Figure 12. A schematic preference tree for the choice between shades of gray.

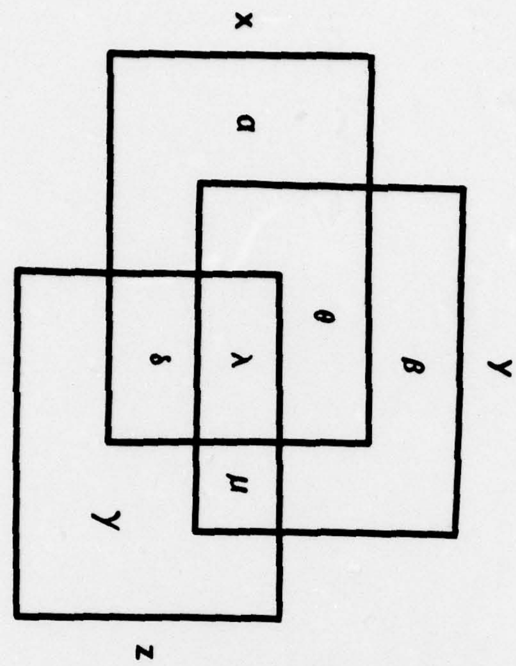


Figure 1. Schematic representation of three alternatives.

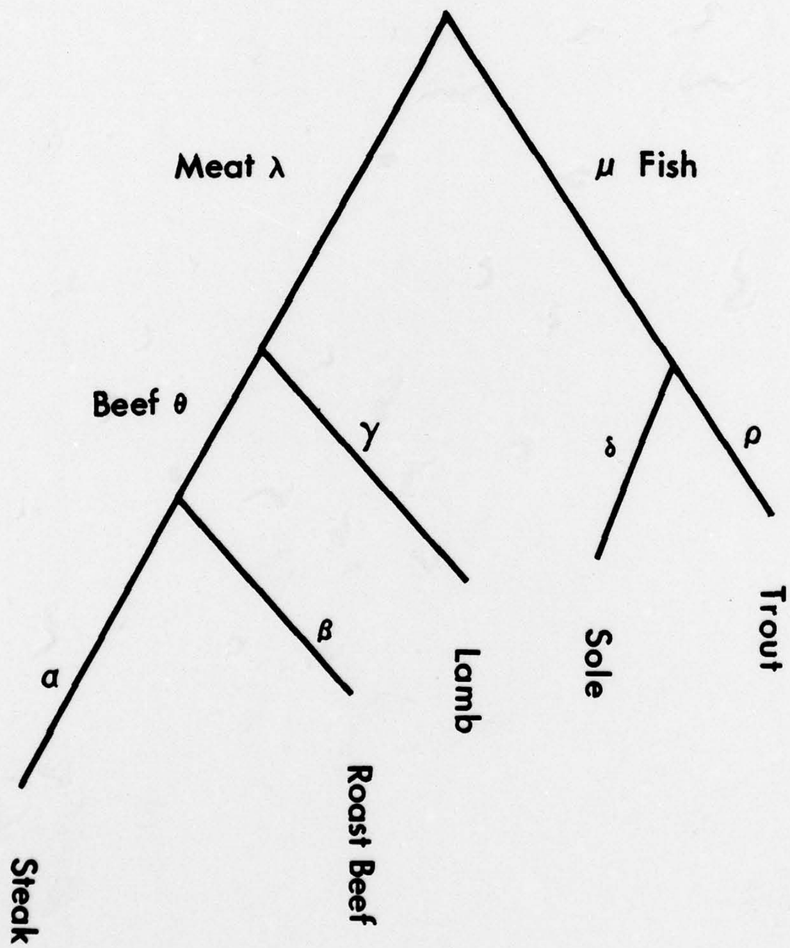


Figure 2. Tree representation of the choice among entrees.

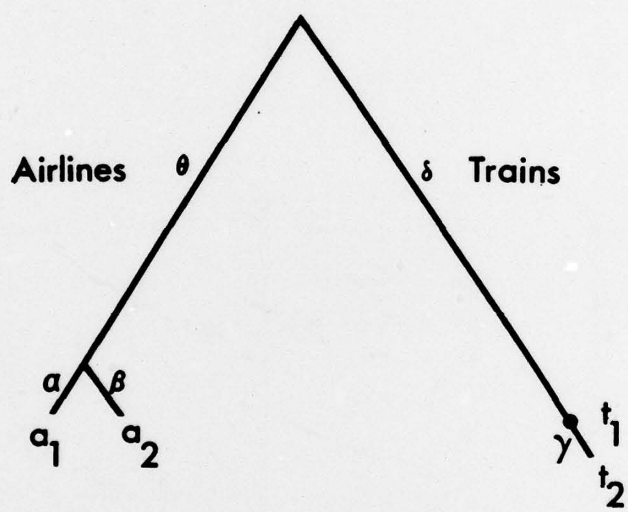


Figure 4. A preference tree for the choice among modes of transportation.

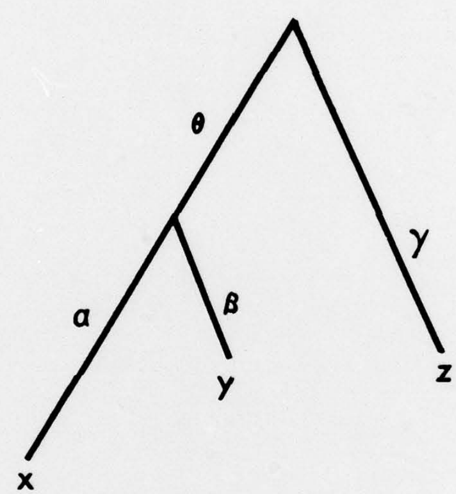


Figure 5. A preference tree for three alternatives.

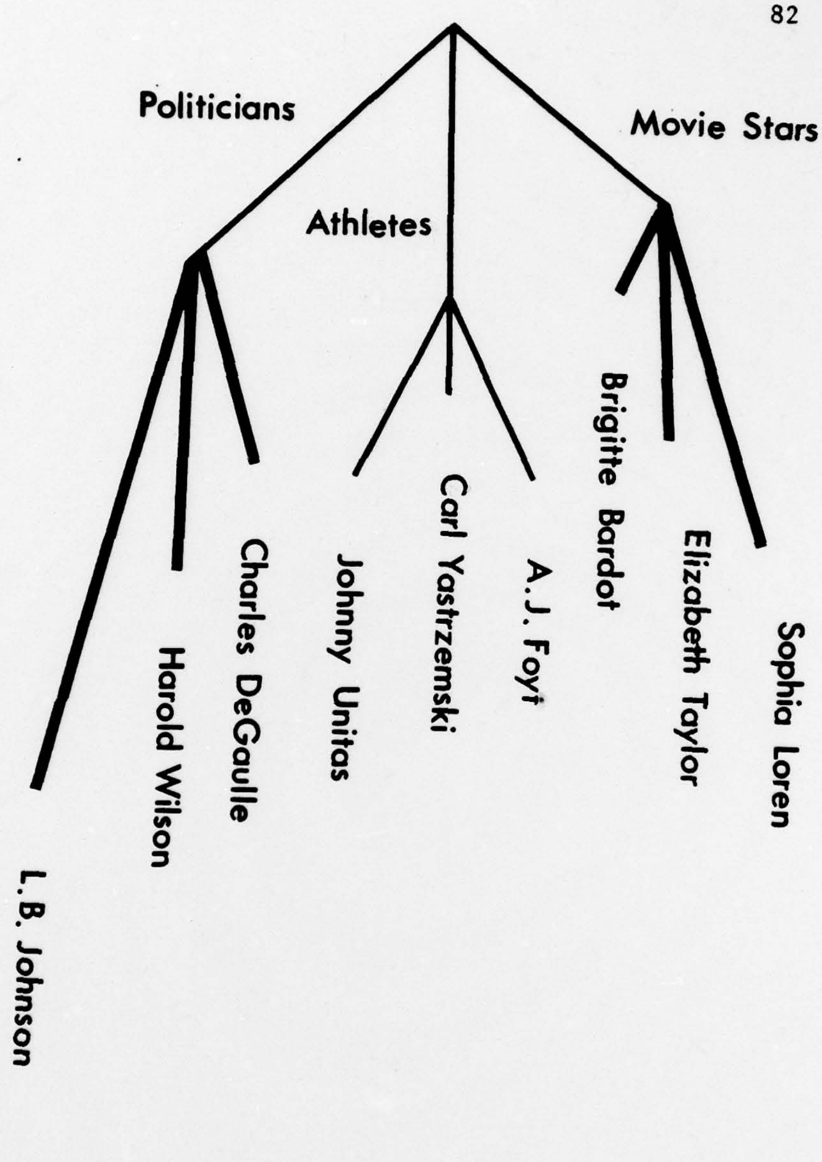


Figure 7. Preference tree for choice among celbrities.

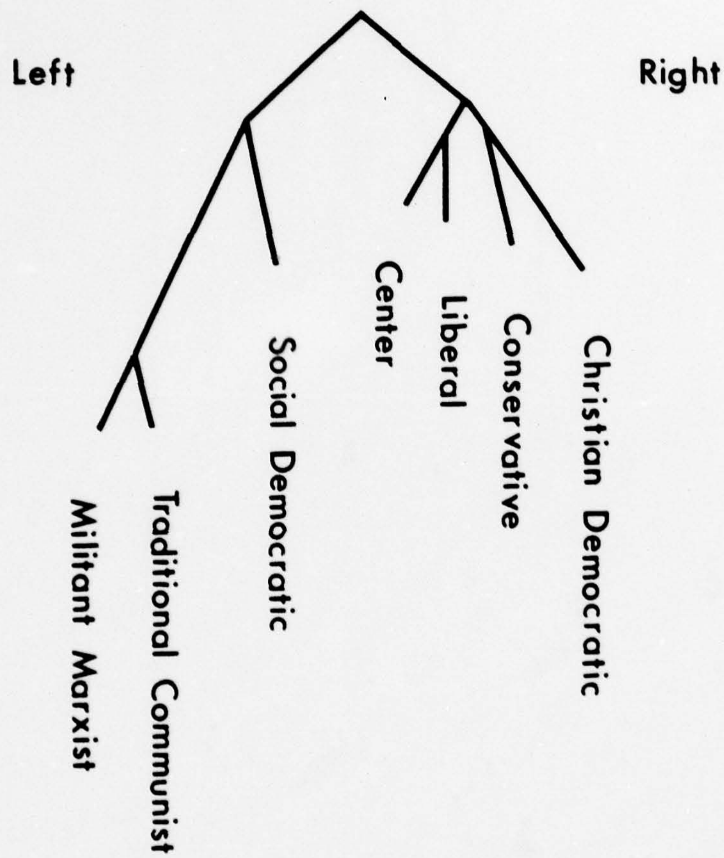


Figure 8. Additive tree (ADDTREE) representation of the similarities between Swedish political parties.

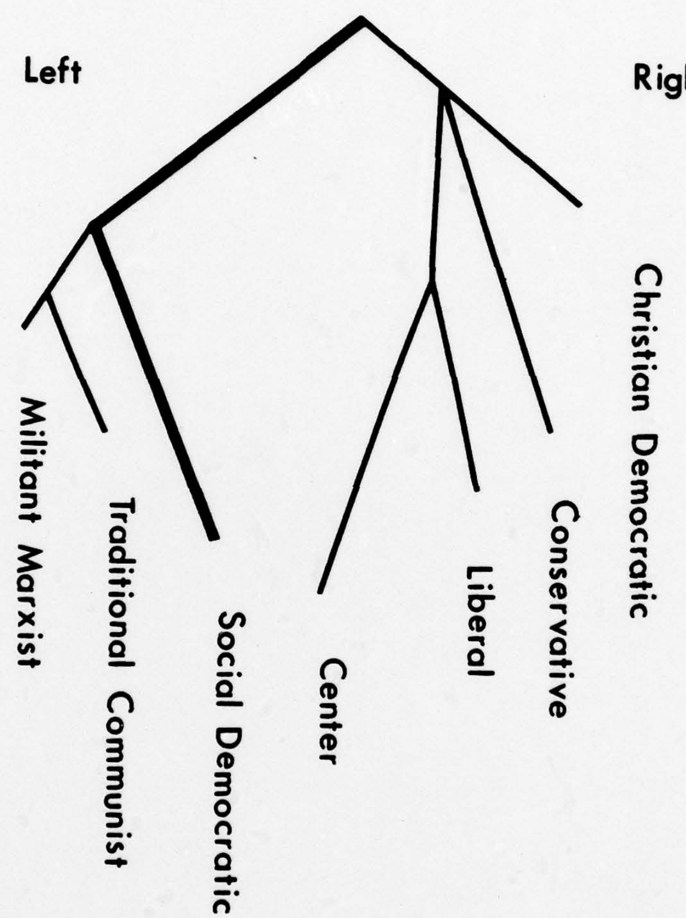


Figure 9. Preference tree for choice among Swedish political parties.

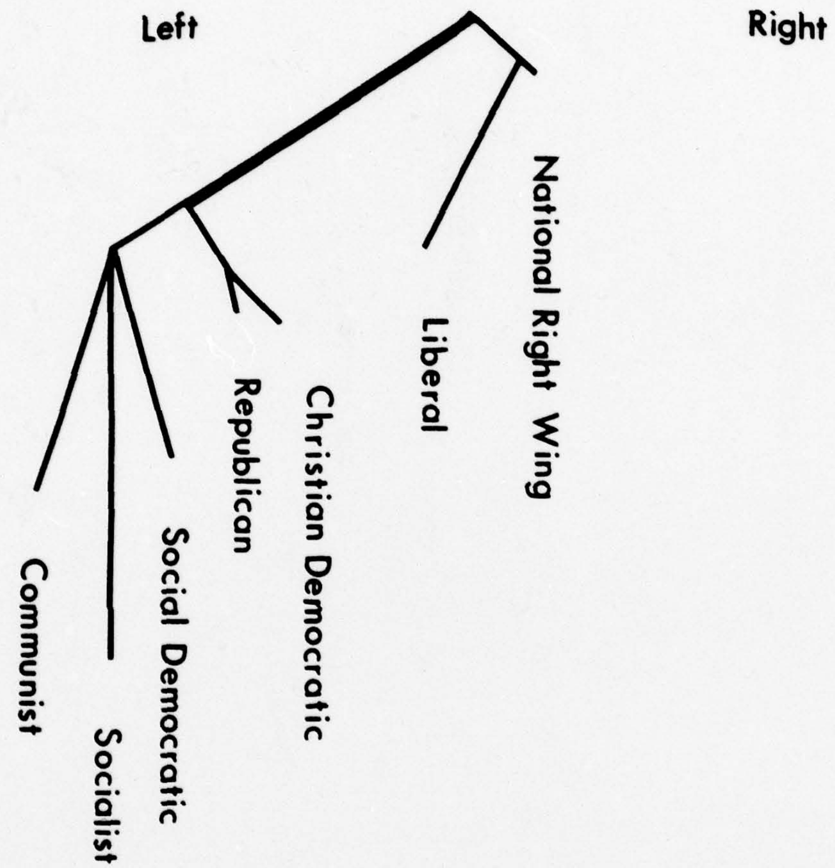


Figure 10. Preference tree for choice among Italian political parties.

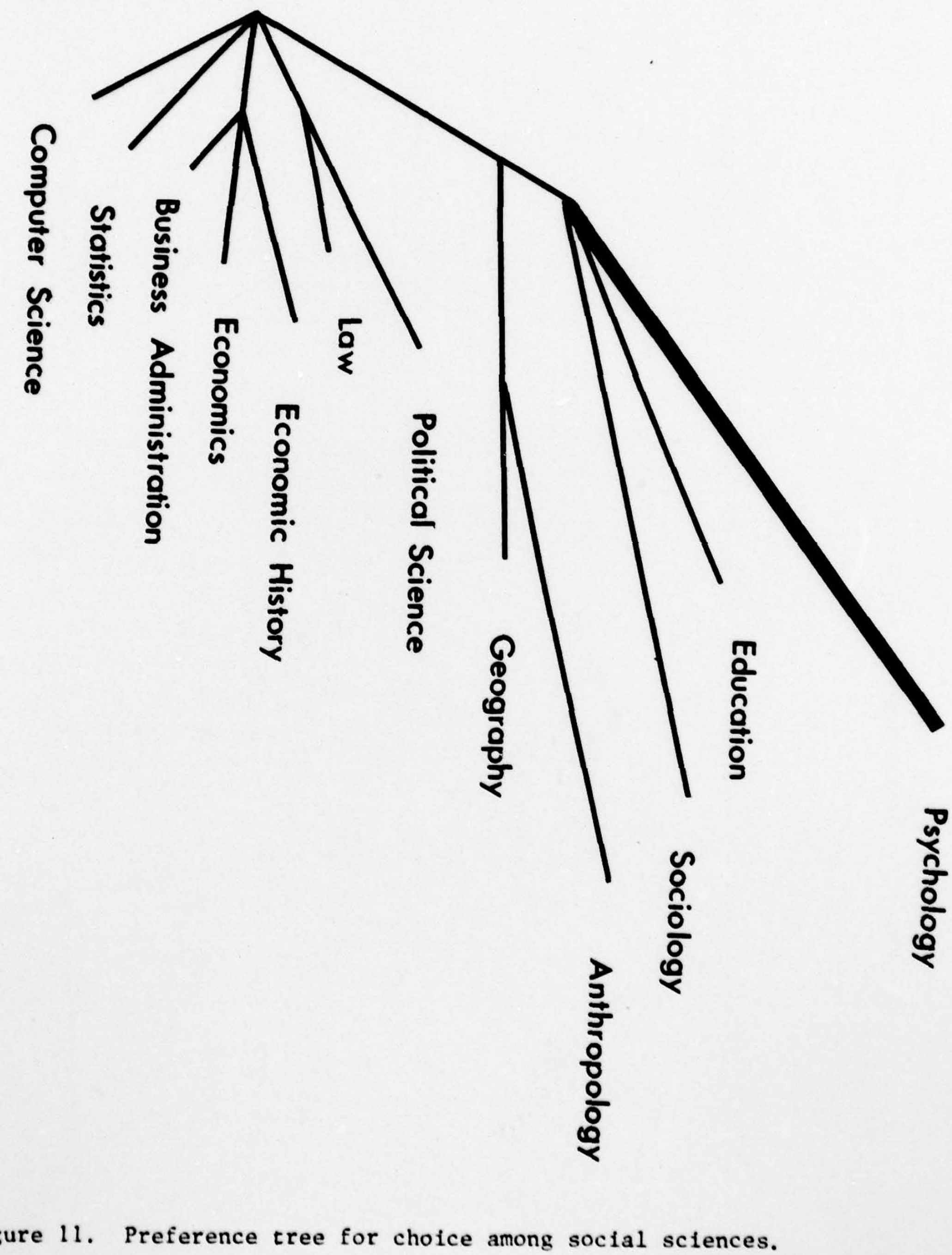


Figure 11. Preference tree for choice among social sciences.

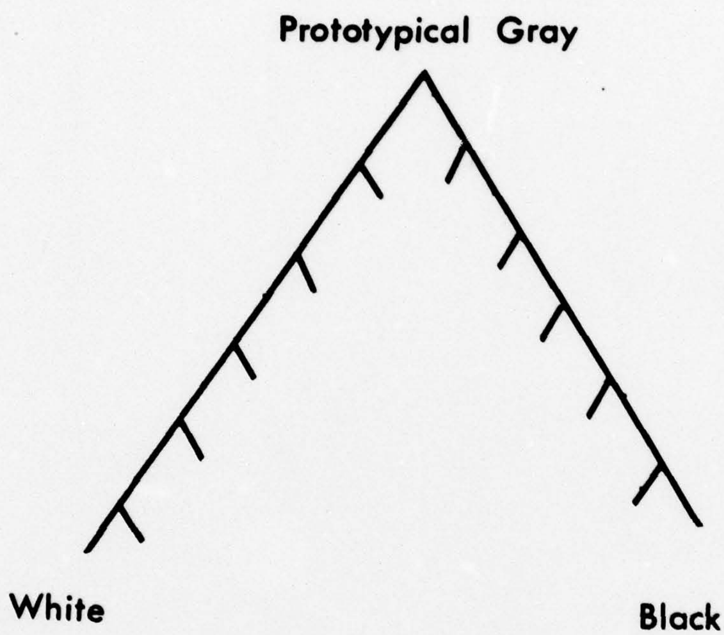


Figure 12. A schematic preference tree for the choice between shades of gray.

DISTRIBUTION LIST

OSD

CDR Paul R. Chatelier
Military Assistant for Training and Personnel Technology
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Department of the Navy

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217 (5 cys)

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Director
Statistics and Probability Program
Code 436
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Commanding Officer
ONR Branch Office
ATTN: Dr. J. Lester

Department of the Navy

Building 114, Section D
666 Summer Street
Boston, Massachusetts 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
536 South Clark Street
Chicago, Illinois 60605

Commanding Officer
ONR Branch Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, California 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, California 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Dr. Robert Smith
Office of the Chief of Naval Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, Florida 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-OOT
Orlando, Florida 32813

Department of the Navy

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, California 94940

Dean of Research Administration
Naval Postgraduate School
Monterey, California 94940

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 98T24
Washington, D.C. 20360

Scientific Advisor to DCNO (MPT)
OP 01T (Dr. Marshall)
Washington, D.C. 20370

CDR Thomas Berghage
Naval Health Research Center
San Diego, California 92152

Navy Personnel Research and
Development Center
Manned Systems Design, Code 311
San Diego, California 92152

Navy Personnel Reseach and
Research Center
Code 305
San Diego, California 92152

Navy Personnel Research and
Development Center
Management Support Department
Code 210
San Diego, California 92152

CDR P.M. Curran
Code 604
Human Factors Engineering Division
Naval Air Development Center
Warminster, Pennsylvania 18974

Mr. Ronald A. Erickson
Human Factors Branch
Code 3175
Naval Weapons Center
China Lake, California 93555

Dean of Academic Departments
U.S. Naval Academy
Annapolis, Maryland 21402

Department of the Navy

LCDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, California 93940

Mr. Merlin Malehorn
Office of the Chief of Naval
Operations (OP 102X)
Washington, D.C. 20350

Department of the Army

Mr. J. Barber
HQS, Department of the Army
DAPE-PBR
Washington, D.C. 20546

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Dr. Edgar M. Johnson
Organizations and Systems Research
Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, Virginia 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, Maryland 21005

Department of the Air Force

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, Ohio 45433

Department of the Air Force

Air University Library
Maxwell Air Force Base, Alabama 36112

Foreign Addressees

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
England

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
Federal Republic of Germany

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
Establishment
Teddington, Middlesex TW11 OLN
England

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
Canada

Dr. A.D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
England

Other Government Agencies

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, Virginia 22314 (12 cys)

Dr. Stephen J. Andriole
Director, Cybernetics Technology
Office
Defense Advanced Research Projects
Agency
1400 Wilson Blvd
Arlington, Virginia 22209

Other Government Agencies

Dr. Judith Daly
Cybernetics Technology Office
Defense Advanced Research Projects
Office
1400 Wilson Blvd
Arlington, Virginia 22209

Other Organizations

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, NW
Washingtion, D.C. 20036

Dr. Ward Edwards
Director, Social Science Research
Institute
University of Southern California
Los Angeles, California 90007

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, Oklahoma 73069

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Room 201
Boulder, Colorado 80309

Dr. Ronald Howard
Department of Engineering-Economic
Systems
Stanford University
Stanford, California 94305

Dr. William Howell
Department of Psychology
Rice University
Houston, Texas 77001

Human Resources Research Office
300 N. Washington Street
Alexandria, Virginia 22314

Professor Douglas E. Hunter
Defense Intelligence School
Washington, D.C. 20374

Other Organizations

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, NW
Washington, D.C. 20036 (3 cys)

Dr. Clinton Kelly
Decisions and Designs, Inc.
8400 Westpark Drive, Suite 600
P.O. Box 907
McLean, Virginia 22101

Dr. Robert Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, California 93017

Dr. Gary McClelland
Institute of Behavioral Sciences
University of Colorado
Boulder, Colorado 80309

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Menlo Park, California 94025

Dr. Jesse Orlansky
Institute for Defence Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

Professor Judea Pearl
Engineering Systems Department
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, California 90024

Professor Howard Raiffa
Graduate School of Business
  Administration
Harvard University
Soldiers Field Road
Boston, Massachusetts 02163

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, Oregon 97401

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, California 91364